EFFORT OF ODYSSEY OF THE MIND CREATIVE

PROBLEM SOLVING TEAMS: EFFECTS ON

CREATIVITY, CREATIVE SELF-CONCEPT,

LOCUS OF CONTROL AND GENERAL

SELF-CONCEPT IN GIFTED

CHILDREN

By

ANNE S. FISHKIN

Bachelor of Arts
Shimer College
Mt. Carroll, Illinois
1958

Master of Arts
University of Colorado
Boulder, Colorado
1970

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1989

EFFORT OF ODYSSEY OF THE MIND CREATIVE

PROBLEM SOLVING TEAMS: EFFECTS ON

CREATIVITY, CREATIVE SELF-CONCEPT,

LOCUS OF CONTROL AND GENERAL

SELF-CONCEPT IN GIFTED

CHILDREN

Thesis Approved:

Kay S. Bull
Thesis Adviser

J. Barbara Wilkinson

Sharon E. Scholten

Paul G. Warden

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to her friends, family, and colleagues for their assistance.

I would like to thank the members of my committee for their availability and their thoughtful questions. My chairman, Professor Kay Bull, was incredibly generous with his time. His uncompromising quest for accuracy and for clarity of thought increased the challenge already inherent in this complex study. Also, three members of the faculty of Applied Behavioral Studies in Education, Dale Fuqua, Paul Miller, and Michael Warner, although not members of my committee, were most gracious with their time as consultants on various thorny statistical questions.

I appreciate my supervisor, Dr. Laura Beckham, for her permission to conduct this study in the Edmond Public Schools and for her helpful contributions. I am indebted to my colleagues, Kathleen Kaufman and Barbara Sharp, for their willingness to devote their time to this study. Without the cooperation graciously given by the parents of the children in the Edmond Enrichment Program this study could not have been completed.

I greatly appreciate the expert consultation of Drs. John Kauffman and Oliver Anderhalter of Scholastic Testing Service, and their generosity in providing the scoring staff to conduct the two inter-rater reliability studies. I feel honored by the kind words of encouragement from Dr. E. Paul Torrance, and I am grateful for the opportunity to consult with Dr. Orlow Ball during the developmental stages of the verbal streamlined scoring for the TTCT.

I am indebted to Lori Griffin for her cheerful availability and expert skills in typing this manuscript. I wish to thank Sarah Junkin for her time and her careful reading of the manuscript.

In order to devote the concentration and time to bring a study of this magnitude to completion, it often has been necessary for me to become reclusive. To Kathleen, Pat, Marie, and Hugh: your friendship and assistance have been an invaluable source of renewal and strength.

To my daughters, Ruth Ellen and Amy Lynne: you are the inspiration for my dedication to nurturing creativity and effort in gifted children. That inspiration is renewed as I see you, and my son-in-law, George, following your own courses of intellectual and emotional growth. It is with pride and love that I acknowledge your encouragement.

Over the past several years, I have, like it or not, developed a personal relationship with this research. I am grateful for the growth I have experienced in my own sense of creativity during the course of this work. The essence of this thesis applies to its author as well; personal growth is very likely to be associated with high effort in creative enterprise.

TABLE OF CONTENTS

Chapter | Page

Chapter Page

Chapter | Page

Chapter | Page

LIST OF TABLES

List of Tables in Appendix P

Table Page

Table		Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO THE STUDY

Creativity and Mental Health

Creativity (Eberle & Stanish, 1980; Sund & Carin, 1978; Torrance, 1979) and positive self-concept (Sears & Sherman, 1964) are critical characteristics which must be developed in order to actualize potential abilities. Enhancement of creativity and of self-esteem are important goals in educating gifted children (Clark, 1983; Renzulli, 1977) as it has been suggested that learning to live creatively will contribute to mental health (Bull, 1978; Maker, 1982) and/or human fulfillment (Davis, 1983; Parnes, 1967/1972, 1975; Torrance, 1978/1984).

Training for creativity, which includes learning a problem solving process that aims at achieving a unique product and increasing the creative abilities of the student, will help emotional growth and enhance emotional well-being (Koberg & Bagnall, 1976). Because creative acts require integration of complex phenomena, they enable the person to experience essential and direct qualities in feelings, thoughts, and actions. People who are aware of and comprehend broad aspects of life are psychologically healthy (Barron, 1963). In addition, increased feelings of competence as a focused inquirer (Davidman, 1982) and a problem solver (Torrance, 1978/1984) contribute to mental health. Schubert and Biondi (1977) summarized several writers who suggested that a strong relationship must exist between creativity and high self-esteem. The creative process requires that creators trust their own powers of thinking and judgment in interpreting that their improved or alternate solutions are worthy (Schubert & Biondi, 1977). Thus, it is seen that a belief in one's own ability to succeed is a necessary trait to becoming more creative. It is also seen that successful experiences in creative enterprise contribute to affective as well as cognitive growth, that is, to emotional well-being.

The focus of this study is to examine a method of creativity training, team creative problem solving, in relation to measured creativity and to aspects of emotion, specifically general self-concept, creative self-concept, and locus of control (see definitions, pp. 12-18) in gifted children. Creative problem solving (CPS) is considered to provide an essential process and skill for the gifted (Feldhusen & Treffinger, 1985). It strengthens the imagination (Parnes, 1975), divergent thinking, and analytic and evaluative skills in a very clear and powerful manner that is applicable to many situations (Parnes, Noller & Biondi, 1977).

The CPS method of creativity training has been shown to increase cognitive creativity (Rose & Lin, 1984; Torrance, 1972a). "Creative problem solving . . . becomes an instructional method for assisting children to become increasingly resourceful, self-sufficient and productive" (Eberle & Stanish, 1980, p. 12).

Significance of the Study

Two popular programs were recently designed to use CPS processes in team competition of young people. The Future Problem Solving Program has been designed and advocated (Crabbe, 1982, 1985; Torrance, Torrance & Crabbe, 1980-1981) as a particularly effective and motivating system of teaching CPS. In 1978, Gourley and Micklus instituted Odyssey of the Mind, first known as Olympics of the Mind. Odyssey of the Mind (OM) is a creativity group competitive process which fosters the growth of creativity and team problem solving skills among highly able and creative young people (Bull & Fishkin, 1987; Micklus & Micklus, 1987). Experience in the CPS process on an OM team was reported by students, teachers, and coaches to have benefited student problem solving skills, flexibility, and risk-taking (Harrington, 1984). An evaluation study of an elementary enrichment program for gifted students concluded that OM was successful and was experienced to be the most challenging and satisfying of the activities in the program by many of the students as nearly 70% of the seventh grade graduates of the program mentioned OM as the activity which gave them the "greatest feeling of accomplishment, success, and satisfaction" (Miller, 1983, p. 50) while only 12% listed OM as yielding the least sense of accomplishment.

In a rationale for the Future Problem Solving Program, Torrance (1985) stated that "gifted children . . . frequently lack skills of interdependence even more than do children of average ability" (p. 3). Interdependence is encouraged when children work in cooperative learning groups, particularly when using the CPS process (Riley & Parr, 1988). In order to benefit from appropriate group experiences, the gifted must have time to be grouped together for some portion of the school day (Parker, 1983). Training in group interaction skills is an important goal in education of gifted students (Maker, 1982). Interaction skills include competence in how to be a

> . . . good follower, how to ask for help, how to help others feel important and valued, and how to improve the functioning of a group without being the group leader. They can learn to become 'enablers,' delegating and assisting rather than commanding while maintaining a sense of control (Webb, Meckstroth & Tolan, 1982, p. 153).

Training in these interaction skills, particularly through team CPS then, is considered to aid affective development in gifted children, and therefore contribute to their emotional health.

Skill in CPS tasks was found to relate to affective measures (Houtz, Rosenfield & Tetenbaum, 1978). When variability due to school achievement was controlled, locus of control and self-esteem showed a significant relationship to performance on tasks of creative thinking and problem solving. Houtz et al.

found that considerable individual variation occurred within the creative thinking and problem solving tasks which suggested that gifted children need training in these areas. It has been demonstrated that there is a relationship between affect and skill in CPS (Houtz, et al., 1978; Sherman, 1977) and studies of creative adults show a relationship between high self-esteem and creativity (MacKinnon, 1968; Stein, 1968). There is, however, little in the documented literature to show that affective measures change with creativity training (see discussion in Chapter II).

Several recent studies have attempted to measure affective changes in children who were exposed to various creativity training methods. These studies failed to find changes when affect was measured by a general scale of self-concept, the Piers-Harris Children's Self-Concept Scale (Fults, 1981; Jones, 1983; Kolloff & Feldhusen, 1984; Tadlock, 1981). An examination of the Piers-Harris Scale (Piers, 1977) shows that it asks few questions about self-concept in relation to creativity.

When self-concept was measured as a general, global construct, it was found not to be related to objective measures of creativity (Wright, Fox & Noppe, 1975). However, Wright et al. identified a measure of creative self-concept as an element in the assessment of self and this creative self-concept showed a relationship to scores on a test of cognitive creativity, subtests of the Torrance Tests of Creative Thinking (Torrance, 1974a, 1974b). Therefore, in a demonstration testing that CPS can contribute to student affective development, a measure of creative self-concept ought to be more sensitive than a measure of general self-concept.

Training in group interaction skills and in CPS, then, is highly recommended as these are activities which are appropriate to differentiated education for gifted students (Clark, 1983; Davis & Rimm, 1985; Maker, 1982). Research findings of the benefits of group interaction on creativity with various populations are generally supportive (Parnes, 1967/1972; Torrance, 1972a) but some are conflicting (Amabile, 1983) and those on affect are very sparse.

A recent study of five OM teams showed that meaningful differences between the teams may in part be explained by examining characteristics of the children and their coaches (Cohen, 1987). Perhaps if the amount of energy individuals invest in their work during creativity training were to be considered, the relationships would become clearer. Observation of gifted children in class activities and on OM teams by this author supports the view that most of the children thrive on these activities. Some, however, are much less responsive to interactive aspects of the enrichment program and a few prefer to withdraw from the program. These observed differences seem to be in part related to the amount of energy or effort children expend on the program activities.

In examining the relationship of effort and creativity, it has been noted that expenditure of effort appears to be essential to completion of work inspired by creative ideas. Effort associated with creativity has been described by Torrance (1979) as perseverance and emotional involvement and by Renzulli (1978) as task persistence or task commitment. Creative achievement occurs when an individual displays the motivation (commitment), possesses the creative skill, and behaves creatively (Torrance, 1979).

The construct of self-efficacy relates decisions to undertake and persevere at a behavior to expectations of success or failure in a given situation (Schack, 1987). An instrument to measure self-efficacy was developed by Starko (1988) and found to be significantly related to the number of original products created by students. Burns (1988) and Schack (1987) found significant relationships of self-efficacy measures in prediction equations of original, real (Type III) investigations (Renzulli, 1977).

This researcher has been unable to locate research literature specifically investigating the effects of effort or levels of participation in changing creative or affective behaviors of gifted children when they work together as a team on a CPS task. There is, then, a need to conduct a study in which appropriate measures are used to assess change in affect as well as creativity in gifted children when they are trained in CPS. There is especially a need to investigate the effect of team CPS training in gifted cihldren.

Purpose of the Study

There are two major thrusts of the present study: 1) to determine if behavioral changes in creativity and in affect can be reflected as effects of effort in team CPS; and 2) to provide construct validation for self-concept and locus of control measures that are specific to creative behavior in children. More specifically, the primary focus of this study is to test the effect of effort on an OM CPS team on a gifted student's creativity, creative self-concept, locus of control, and general self-concept. The intent of the study is to investigate the effects over time for elementary gifted children in a pull-out enrichment program in relation to the differential effects on the dependent variables of three levels of effort in OM, a voluntary team CPS program available to enrichment students.

The secondary focus of the study is to test the sensitivity of the creative self-concept and locus of control measures to OM CPS team training. The intent is to provide these measures of affect which are logically appropriate to the study of behavioral changes associated with creativity training and to discover their relationship to one another, to creativity training, and to a measure of general self-concept selected for its appropriateness to the study of creativity.

Justification for the Study

Need for Research on Team Creative Problem Solving and Gifted Children

There has been recent discussion of the relationship between self-concept, creativity, and CPS. Several studies investigated self-concept and other affective variables, locus of control, and creativity in CPS tasks when gifted children were working individually (Beck, 1979; Houtz, Montgomery, Kirkpatrick & Feldhusen, 1979; Sherman, 1977). These studies investigated the inter-relationships of the variables but did not utilize repeated-measures (pre- and posttesting) designs to assess effects of training in CPS.

Authors of non-experimental articles may be found who assert that group CPS experiences benefit self-concept and leadership skills in gifted students (Crabbe, 1982; Parker, 1983) but they provided no data. Children in regular classes worked cooperatively within teams on non-CPS problems and demonstrated positive affective changes including growth in leadership skills (Slavin & Karweit, 1981). When students worked in cooperative academic teams they showed higher levels of self-esteem, liking of others, and feeling liked by others (Slavin, 1986). With the exception of a study by Foster (1981) investigating creativity of cooperative learning groups in elementary science problem solving tasks, and studies by Cohen (1987) and by Hendrickson (1985/1986) on team CPS in gifted children, there is a scarcity of experimental research investigating influences on creativity in groups or on teams.

> There is little relevant theory {on social psychology of creativity}, there is only a small research literature on the effect of specific social and environmental influences on creativity and, more importantly, there are virtually no experimental studies of the effects of such influences (Amabile, 1983, p. 3).

Studies of the effectiveness of CPS have demonstrated that CPS increases creativity in college students (Mansfield, Busse & Krepelka, 1978; Rose & Lin,

1984) and in children (Maker, 1982; Torrance, 1972a). However, "no research has concentrated on the effectiveness of CPS with only . . . gifted" (Maker, 1982, p. 201) children other than the recent study by Hendrickson (1985/1986). With the exception of the studies by Cohen (1987) and by Hendrickson (1985/1986), this review of the literature has failed to locate research to substantiate the effect on creativity of group CPS as a treatment. An ERIC search (April, 1988) yielded only three researchers who reported data gathered on students in work on Future Problem Solving teams (Hendrickson, 1985/1986; Tallent, 1985; Torrance & Mourad, 1978).

In 1985, Dr. E. Paul Torrance (personal communication, July, 1985) (see Appendix A, Documentation of Personal Communications) confirmed that no studies were known to him which measure effects of Future Problem Solving in relation to changes in student creativity. No studies of changes in children's creativity in relation to work on a Future Problem Solving team were known to the leadership of the Future Problem Solving Program (A. Crabbe, March 12, 1987, July 7, 1988; or G. Shewach, March 12, 1987; personal communications, see Appendix A). The study by Hendrickson (1985/1986) studied the effect of Future Problem Solving and other CPS on gifted students, but did not use a pre- and posttest design. No studies have been completed yet which investigate the effect of OM training on children's creativity when they participate in an OM team (C. S. Micklus, personal communication, November 10, 1987, see Appendix A).

Thus, there is a need to study the effects of team CPS on creativity in children, and on aspects of their self-concept relevant to creative behavior, more specifically, creative self-concept and locus of control. Because CPS is recommended and widely used in gifted programs (Crabbe, 1985; Feldhusen & Treffinger, 1985; Micklus & Micklus, 1987; Torrance, 1979), there is especially a need to study effectiveness of group or team CPS in gifted children.

Need for Multivariate Analysis in Creativity and Self-Concept Research

Creativity is a complex phenomenon which is expressed by a multitude of behaviors (see pp. 25-31, Chapter II). It is therefore appropriate in the study of creativity to examine several dependent variables as measures of various aspects of creative behavior. The appropriate method of analysis for a study of several related behaviors is the use of multivariate statistical procedures (see definition, p. 16, and Appendix B, Procedural Issues in Multivariate Statistics).

> Because of the number and extent of aptitude factors involved in creative talent, it is unlikely that any small, relatively arbitrary selection of tests will predict well a complex multidimensional criterion of creative behavior. This suggests, in addition to the need for broadening the selection for test tasks, the need to utilize complex multivariate statistical procedures rather than simple bivariate correlational procedures (Treffinger & Poggio, 1972, p. 257).

The measurement of cognitive creativity using divergent thinking tests generally yields multiple measurements such as fluency, flexibility, and originality as on the Torrance Tests of Creative Thinking (TTCT) (Torrance, 1974a, 1974b). Several studies, such as those in basic movement (Tadlock, 1981) or creative dramatics (Bennett, 1982; Myerson, 1981) utilized the TTCT and two or three measures aimed at assessing changes in student attitude towards self, such as self-esteem. Some have used multiple affective measures, such as self-esteem, locus of control, and tolerance of ambiguity (Houtz, Rosenfield & Tetenbaum, 1978; Tetenbaum & Houtz, 1978). All of the foregoing studies used analysis of variance (ANOVA) to analyze the data. The use of multiple ANOVAs to examine effects on several dependent variables that are related is inappropriate because it increases the likelihood of finding false positive or chance results. In one study (Bennett, 1982), multiple t-tests and multiple Pearson product moment correlations were utilized and an increase in scores on the Piers-Harris was found after the treatment. The study, however, could have inflated significance levels due to chance occurrences. Problems with inappropriate data analyses are considered in more detail in Appendix B (Procedural Issues of Multivariate Statistics). The use of multivariate statistics provides the appropriate procedures (see definition p. 16 and Appendix B) for analysis of multiple related dependent variables (Tabachnick & Fidell, 1983; Treffinger & Poggio, 1972).

A few multiple regression studies (Kogan & Pankove, 1974; Schack, 1987; Starko, 1988) have appeared in the creativity literature. Such studies are more appropriate because multiple variables can be examined in combination to investigate prediction of creative behavior. Other studies have used multivariate analysis of variance (MANOVA) procedures to test effects of several related variables as a single multivariate construct. A MANOVA was used to examine effects on the construct of self-concept and creative thinking of an enrichment program which included some experiences in group and individual CPS (Kolloff & Feldhusen, 1984).

The present study investigated several similar variables as Kolloff and Feldhusen (1984), but differs in using two multivariate constructs. To defend use of a multivariate analysis, the variables must belong together not only statistically, correlated greater than .3, but also conceptually (Finn & Mattsson, 1978). There is a conceptual difference between a product measure such as responses to the TTCT or Similes and a personality measure such as the self-report inventories forming the affective construct (Isaac & Michael, 1981), and, therefore, two constructs are necessary.

In this study, the creativity measures are the figural and verbal batteries of the TTCT (Torrance, 1974a; Torrance & Ball, 1984) and Similes (Schaefer, 1971). The personality or affective measures used in this study are: general self-concept, Sears Self-Concept Inventory (Sears) (Sears, 1975); locus of control, Intellectual Achievement Responsibility Scale (IAR) (Crandall, Katkovsky &

Crandall, 1965); and creative self-concept, How Many Ideas? (Ideas) (Fishkin, 1987a, 1987b).

Need for Creativity Research Which Uses Sensitive and Appropriate Affective Instrumentation

Many studies fail to find significant differences when the measuring instruments used are trait measures or are for some other reason insensitive to change. In some cases the measures do not contain content related to the treatment and, therefore, they cannot reflect changes associated with the treatment even if changes do occur.

In reviewing recent research of effects on affect, several studies were found which used the Piers-Harris Children's Self-Concept Scale and which failed to find the hypothesized positive effects on self-concept of the creativity training program (Myerson, 1981; Tadlock, 1981) or an enrichment program with a major emphasis in creativity (Fults, 1981; Kolloff & Feldhusen, 1984). In the studies by Fults (1981), Myerson (1981), and Tadlock (1981) the use of the Piers-Harris Children's Self-Concept Scale did not yield increases in children's general self-concept associated with various kinds of creativity training. Perhaps these training programs may indeed not have been effective in relation to student affect. Or perhaps, the programs might have been found to be effective if a scale of creative self-concept were used, but the Piers-Harris (Piers, 1977) does not include questions with specific content appropriate to creativity training. Or perhaps the training program might have some effects on aspects of affect related to but distinct from general self-concept, such as locus of control, preference for seeking sensation or risk (Davis & Subkoviak, 1975), or tolerance for ambiguity (Tetenbaum & Houtz, 1978). The question of whether the creativity program was indeed effective, but in need of measurement by a more specific attitude toward self (Shavelson & Bolus, 1982) as in creative self-concept (Wright et al., 1975) does not seem to be adequately addressed in their articles. It is discussed in Chapter II in greater depth.

Definition of Terms

Affective domain. The affective domain encompassess characteristics of learning and curriculum in which the primary focus is on feelings and/or interpersonal behavior (Krathwohl, Bloom & Masia, 1964). Included in this broad category are student self-concept, self-esteem, attitudes, values, and self-motivated learning characteristics such as locus of control (Clark, 1983; Davis & Rimm, 1985).

Creative problem solving: Creative problem solving (CPS) is a multi-stage process which encourages generation of many ideas in a brainstorming session. A CPS model was originated by Osborn (1957) and further developed by Parnes and his associates (Parnes, et al., 1977). It differs from the usual problem solving approaches in its recognition of and emphasis on creativity as a vital part of the process (Maker, 1982). The person who uses CPS learns to separate the divergent and convergent thought processes; and to defer judgment when involved in the production of ideas. The idea production phase is alternated with analysis and judgment phases of those ideas. The CPS process is best used with small groups in order to maximize interaction among the members of the group (Firestien & Treffinger, 1983). The teacher acts primarily as a facilitator of the group process with a recorder and a "back up" recorder often chosen to write down ideas generated by the group (Firestien & Treffinger).

In this study, the CPS model will be more specifically addressed to team CPS using Future Problem Solving techniques and to OM. In Future Problem Solving the CPS process follows the formal Osborn-Parnes model quite closely, but the responsibilities are shared by team members and the teacher takes an indirect role as coach to facilitate the process. Rather than appointing a recorder, Future Problem Solving team members are asked to write down their own ideas during brainstorming phases and also to record the ideas arrived at by group consensus.

In OM, likewise, each team member is encouraged to take responsibility for his or her contributions to the team, and the coach may only serve as a facilitator, for the ideas and work must come from the team. In OM teams, because the outcomes of the team's work are of a hands-on, performance nature, the ideas of the brainstorming phases frequently may not be recorded by the coach or team members. Coaches of OM teams for this study were briefly introduced to the CPS model developed by Bull (Bull & Fishkin, 1987). Aspects of OM training which most closely follow the Osborn-Parnes CPS process are incorporated in the practice sessions for brainstorming of the spontaneous problems. The coaches of OM in this study were parent volunteers who received only two hours of training in OM techniques (see Chapter III and Appendix C) and only a small portion of that training was devoted to CPS. Therefore, the findings of this study may not be directly comparable to the outcomes of research studies of the Osborn-Parnes process of CPS.

Creativity: Many definitions have been proposed for creativity, focusing primarily on a process, a product, a personality, or an environmental emphasis (Arieti, 1976). The creativity definition proposed for research purposes by Torrance (1974b) is accepted as a satisfactory definition to describe the process involved in the team CPS activities, OM and Future Problem Solving. Torrance defined the process of creativity as

> . . . becoming sensitive to problems, deficiencies, gaps in knowledge, missing elements, disharmonies, and so on; identifying the difficulty; searching for solutions, making guesses, or formulating hypotheses about the deficiencies: testing and retesting these hypotheses and possibly modifying and retesting them; and finally communicating the results (Torrance, 1974b, p. 8).

The creative product is the observable outcome of the creative process (Besemer & O'Quin, 1986). It may have physical qualities, or may be a theoretical system (Brogden & Sprecher, 1964, cited in Besemer & Treffinger, 1981). Parnes defined

> . . . creative behavior as that which demonstrates both uniqueness and relevance in its product. The product may be unique and relevant to a group or organization, to society as a whole, or merely to the individual himself. Creativity is thus a function of knowledge, imagination and evaluation . . . Without knowledge, imagination cannot be productive. Without imaginative manipulation, abundant knowledge cannot help us live in a world of change. And without the ability to synthesize, evaluate and develop our ideas, we achieve no effective creativity (Parnes, 1967/1972, pp. 6-7).

For the purpose of this study, which aims to measure changes in observable creative behaviors, a product definition of creativity is necessary. Divergent thinking tests which weigh relevance and uniqueness in the scoring can be considered to be product measures of creativity. The creativity measures used in this study, Similes and the TTCT verbal and figural, give no credit to responses which are irrelevant to the tasks. The operational definition of creativity used in this study is delimited by scores of the three tests of creative thinking. A high score on Similes (Schaefer, 1971) will reflect fluency but is more dependent upon the quality of the responses. High scores on the verbal TTCT battery of activities are reflected in the divergent production measures of fluency, flexibility, and originality where originality is defined as statistical rarity of responses (Torrance, 1974a). High scores on the figural TTCT battery are reflected in the measures of fluency and originality but also in measures of elaboration, abstractness of titles, resistance to premature closure, and a variety of creativity indicators which are summed to a total score of creative strengths. A composite score which reflects the creative strengths and the other five norm-referenced scores, the Creativity Index, is recommended for some users (Torrance & Ball, 1984). These latter measures have been added for the figural measures in order to reflect essential qualities of synthesis and of richness present in creative thinking which are not reflected in the divergent production measures (Torrance, 1979; Torrance & Ball, 1984).

Effort: The degree of participation or energy exerted in a particular task, or a specific area of performance, sometimes called task commitment, is considered to be a critical component in determining the productivity of gifted individuals (Renzulli, 1978; Tannenbaum, 1983; Torrance, 1979; Whitmore, 1980). In this study, effort will be taken to mean how much work someone performs on a task. The amount of work may include time spent on the task, but is unlike persistence which, to this author, is a construct limited to how long someone stays on a task. In this study, effort is operationally defined by peer and self-rating of amount of time and energy spent and work produced on an OM team.

Gifted Children: Public Law 97-35, passed by Congress in 1981, defined gifted and talented children as:

> . . . children who give evidence of high performance capability in areas such as intellectual, creative, artistic, leadership capacity, or specific academic fields, and who require services or activities not ordinarily provided by the school in order to fully develop such capabilities, (Sec. 582) (Clark, 1983, p. 5).

A psychosocial definition of gifted children was proposed by Tannenbaum (1983) to account for the psychosocial aspects of giftedness and the difference between promise and fulfillment of potential.

> Keeping in mind that developed talent exists only in adults, a proposed definition of giftedness in children is that it denotes their potential for becoming critically acclaimed performers or exemplary producers of ideas in spheres of activity that enhance the moral, physical, emotional, social, intellectual, or aesthetic life of humanity (Tannenbaum, 1983, p. 86).

In Oklahoma, "'gifted and talented children' means . . . those identified students who score in the top three percent (3%) on any national standardized test of intellectual ability" (Title 70, 1987, p. 2338). School districts may, in addition to the mentally gifted, elect to serve students who excel in other areas of giftedness. In meeting the state mandate at the elementary level, the school district in this study serves the category of children gifted in intellectual ability. Children gifted in intellectual ability are considered to be children who are high in IQ (Getzels & Jackson, 1962). High intellectual ability, as measured by the Stanford-Binet Intelligence Scale, will usually encompass ability in vocabulary, skillful verbal thinking, abstract thinking including classifications and analogies, capacity of problem solving, understanding of causal relationships, and recall (Clark, 1983). For this study, all children will have individually tested IQs at or above the 97th percentile to be considered gifted.

Locus of Control: The degree to which a person expects reinforcers to be within the realm of personal control quantifies the dimension of internal or external control of reinforcement. It is the

> . . . degree to which the individual perceives that the reward follows from, or is contingent upon, his own behavior or attributes versus the degree to which he feels the reward is controlled by forces outside of himself and may occur independently of his own actions (Rotter, 1966/1982, p. 171).

A person who views positive and/or negative events as related to personal actions is at the internal control end of the continuum, whereas a person who fails to feel personal control for his behavior and who perceives events as unrelated to his own actions is at the external control end of the continuum (Lefcourt, 1966). In this study, locus of control is defined by scores on the Intellectual Achievement Responsibility Questionnaire (IAR) (Crandall, Katkovsky & Crandall, 1965).

Multivariate Statistics: Multivariate statistics provide a method for analyzing the effects of one or more independent variables on several dependent variables in a single analysis. Multivariate statistics are increasing in use in the behavioral sciences and other areas of research (Tabachnick & Fidell, 1983). In order to use multivariate statistics in a research study, the various dependent measures to be considered in the analysis must be related tc conceptually and must bear some correlation to each other; dependent variables must form a global multivariate constr 1978). Further description of multivariate statistics and so are discussed in Appendix B.

Odyssey of the Mind (OM): Odyssey of the Mind was originally conceived as Olympics of the Mind, a team CPS competition program (Gourley and Micklus, 1978). Students who elect to work in an OM team meet under the guidance of a volunteer coach trained in the CPS process. Team members develop a workable solution to their choice of one of five open-ended problems. Teams receive points in three categories of competition: long-term, spontaneous, and style. The OM condition is defined in further detail in Chapter III and Appendix C.

Persistence: Persistence, sometimes referred to as task persistence, is usually operationally defined in the literature as the amount of time an individual will continue working at a specific task in the absence of explicit reinforcement (Gordon, Jones & Short, 1977). Other authors have defined persistence more broadly as willingness to engage in hard work (Franks & Dolan, 1982). To the present author, persistence refers only to how long someone worked and effort refers to how hard they worked.

Self-concept: Self-concept can be defined either in a general way or it can be related to a specific type of performance. The individual's general cognitive picture of the self is complex, and is comprised of the many specific areas of images the person holds, such as physical self, ideal self (aspirations), and real self (comparison with others). Moreover, each of these ways of measuring general self-concept is composed of specific aspects of self-concepts, such as creative self-concept or academic self-concept (Whitmore, 1980). Self-concept differs from self-esteem, which is the measure and feelings of worth, or height of the self-concept (Whitmore, 1980). For this study, the general self-concept will be defined by a reported measure of self-concept, operationally defined by the Sears Self-Concept Inventory (Sears, 1975). The image one holds of the self as a creative person would be measured by a scale of self-concept which has content specific to perceived creative behaviors. A scale of specific creative self-concept was developed by the author for this study. How Many Ideas? (Ideas) (Fishkin, 1987a, 1987b; see Appendix K) is comprised of 10 Likert-like items and is based on the work of Wright et al. (1975).

Team: In the present study, unless otherwise specified, the word "team" is synonymous with an OM team of three to seven children who work under the guidance of a trained parent or volunteer teacher coach, primarily after school. Teams were organized in January and worked together until they were eliminated at a level of competition, local, state, or national.

Teamwork: Teamwork is characteristic of group situations where people have learned the specific skills of cooperation, collaboration (Osborn, 1957), and facilitation of interpersonal processes which are applied to reach a common goal. More specifically, for group CPS, teamwork is characterized by flexibility, enthusiasm, and deferred judgment during the decision-making stages, willingness to keep a problem moving to other team members, and active listening to other members' ideas. Children who are strong in such teamwork skills are likely to display a willingness to try new ideas, to request others to express their ideas, and to complete their fair share of the team's work.

Statement of the Problem

The problem to be investigated is: what is the effect of effort on an OM team on a gifted student's creativity, self-concept, and locus of control? In examining the problem, we will define OM as a form of team CPS and measure effort at three levels (high effort in OM, low effort in OM, or no effort in OM). A composite score for the OM participants consisted of a ranking of the individual student's OM effort by self-report and peer ranking of the amount of effort in OM reported by each member of the team. Those children who chose not to be in OM were used as no effort in OM controls. All children in the study, including the control non-OM group, attended the pull-out enrichment program for gifted children (see Chapter III and Appendix D). Gifted students are defined solely by evidence of mental giftedness as those children who score at the 97th percentile or higher on an individual test of intelligence.

The creative construct was measured by: 1) streamlined scoring of the Torrance Tests of Creative Thinking (TTCT) figural battery which provides six subscale scores (Torrance & Ball, 1984), 2) standard scoring of the TTCT verbal battery which provides three subscale scores (Torrance, 1974a), and 3) Similes, a single score (Schaefer, 1971). The affective construct was measured by: 1) creative self-concept, How Many Ideas? (Ideas) (Fishkin, 1987a, 1987b); 2) an overall self-concept scale, the total score of the Sears Self-Concept Inventory (Sears) (Sears, 1975); and 3) locus of control, the Intellectual Achievement Responsibility Scale (IAR) (Crandall, et al., 1965) with two subscales. These dependent variables were combined to form two multivariate constructs: one for creativity and one for affect. These multivariate analyses establish the relative contributions of the different dependent variables to each other in relation to the independent variables of time and of the OM condition. The variables of prior experience in OM and experience in the enrichment program were tested for their relation to effort in OM.

Statement of Research Hypotheses

There is evidence in the research literature (cited in Chapter II) that training in CPS increases scores on creativity tests. There are conflicting results of studies which examined the effects of CPS training on self-concept or locus of control. Creative self-concept has not yet been tested as a dependent variable affected by CPS training. To date only a preliminary case study approach has examined creativity test scores and general self-concept in OM participants (Cohen, 1987); the effects of OM CPS training have not yet been examined for any of these variables. There are several new factors in this research study: the effect of OM CPS training on creativity variables, the use of multivariate analyses of the TTCT where the figural test is scored by the streamlined procedures, the use of differences in effort as contributing to the levels of the independent variable, and the use of a creative self-concept scale as a dependent variable in the effects of CPS training. There are, therefore, several aspects of this study for which there is no prior evidence to warrant a directional hypothesis.

It was not possible to randomly assign the children to the different OM categories and there were insufficient subjects to permit random selection from a larger population, therefore, this study cannot be categorized as true experimental research. By use of subjects from intact naturally occurring groups, the study is classified as quasi-experimental research which does not provide the rigorous controls necessary to seek causal explanations of research questions (Isaac & Michael, 1981; Tallmadge, 1988). Because there are several previously unresearched aspects of this study and because it is quasi-experimental research, it is appropriate to state the research hypotheses in a non-directional rather than a directional form.

It was hypothesized that there would be significant changes in the creativity construct in all groups of subjects who participated in the gifted program. It was hypothesized that there would be a difference in pre-post scores between students who are experienced and those who are not experienced in the enrichment program. It was likewise hypothesized that there would be a difference in pre-post scores in students with prior experience in OM and students who had no previous experience in OM. It was further hypothesized that there would be a difference in pre-post scores between those students who scored high in their amount of effort in the OM program (high OM condition) and those who scored low in their effort or those who were not on an OM team. It was also hypothesized that there would be differences in creative self-concept, internal locus of control, and in general self-concept, in relation to levels of student effort in OM.

The following hypotheses were tested at the .05 level of significance:

1. There will be a significant interaction on the creativity construct scores with respect to effort in OM and time.

2. There will be a significant interaction on the affective construct scores with respect to effort in OM and time.

3. There will be a significant interaction on the creativity construct scores of students in relation to the amount of their experience in the enrichment program and time.

4. There will be a significant interaction on the affective construct scores of students in relation to the amount of their experience in the enrichment program and time.

5. There will be a significant interaction on the creativity construct scores of students in relation to the amount of their prior experience in OM and time.

6. There will be a significant interaction on the affective construct scores of students in relation to the amount of their prior experience in OM and time.

7. There will be a significant interaction on the creativity construct scores when these scores are compared with respect to effort in OM, experience in the enrichment program, and change in score from pretest to posttest in the experimental year.

8. There will be a significant interaction on the affective construct scores when these scores are compared with respect to effort in OM, experience in the enrichment program, and change in score from pretest to posttest in the experimental year.

9. There will be a significant interaction on the creativity construct scores when these scores are compared with respect to effort in OM, prior experience in OM, and change in score from pretest to posttest in the experimental year.

10. There will be a significant interaction on the affective construct scores when these scores are compared with respect to effort in OM, prior experience in OM, and change in score from pretest to posttest in the experimental year.

11. There will be a significant difference in sensitivity of the scores of the univariate dependent variables of the affective construct to student effort in OM.

12. There will be a significant difference in creative self-concept scores with respect to student effort in OM and time.

Summary

In this chapter a review of the literature revealed the merit of training for creativity and for skills in effective group problem solving in gifted children. The review revealed only two investigations on the effectiveness of team functioning in gifted children who participated in CPS. The need was established for the present study; a multivariate, quasi-experimental investigation of the effects of team CPS in OM in gifted children. The problem was defined, terminology defined, variables were operationally defined, and the hypotheses were stated.

CHAPTER II

REVIEW OF THE LITERATURE

Organization of the Review

In this chapter the major concepts of the study are set forth and examined in light of relevant theoretical and experimental work. The broad field of creativity is briefly reviewed in light of major areas of definition. Issues pertaining to measurement of creativity are reviewed in terms of general issues and issues specific to the creativity instruments used in the study.

Issues pertaining to gifted children are briefly introduced to relate the rationale for creative problem solving, creativity, and affective development to the needs of the gifted. The research on CPS, as the broader area of OM team problem solving, is reviewed in the areas relating to gifted children, to team problem solving in gifted children, and to the effectiveness of CPS as a technique of increasing creativity.

Broad methodological issues which impinge on creativity and affective research studies are briefly analyzed. The need for studies which use multivariate statistics to analyze the data of complex behaviors is presented. The affective constructs investigated in this study, self-concept and expectancy of locus of control of reinforcement, are discussed as broad issues. The need for research using those affective measures which are specific to the content of creativity is discussed for both of these constructs, and the literature relating to the specific measures to be used is reviewed where it relates to gifted children, CPS, and/or creativity measures.

Literature relating to task persistence or effort, and to a social psychological approach to creativity in groups is briefly introduced to aid in understanding the definition of the independent variable as high or low effort of team CPS. The effect of experience in a gifted program is briefly reviewed to establish the need for the two variables which assess relevant previous experience.

In this chapter, then, evidence is presented to justify the different variables as measures of creativity and as measures of behavioral affect as two coherent constructs which are justifiably different from each other. Evidence is also presented to justify the ordering of the various measures within each of the two constructs.

Creativity

Four major perspectives have been offered as a framework to examine creativity: the creative person, including examination of the historical retrospective point of view; the creative process; the role of the environment in cultivation of creativity; and the creative product (Arieti, 1976; Welsh, 1975). These four areas are briefly described and creativity as a construct distinct from intelligence is touched upon. The focus of the study moves then to examination of creative products and issues of measurement to general measures of creative products and to the specific measures, the Torrance Tests of Creative Thinking and Similes.

Major Orientations in a Definition of Creativity

Creativity is a complex and intriguing phenomenon which is ephemeral when writers attempt a definition that reaches broad consensus (Isaksen, Stein, Hills, & Gryskiewicz, 1984; MacKinnon, 1975). One system of classifying theories of creativity was formulated by Gowan in 1972. He described five groups of theories which fit on a continuum from the rational to the psychedelic: 1) cognitive, rational and/or semantic; 2) personality, including environmental; 3) mental health; 4) psychoanalytic, Freudian, and neo-Freudian; and 5) psychedelic. The cognitive (or rational) theorists focus primarily on semantic or verbal phenomena, cognitive abilities and/or gestalt approaches. The semantic theorists may acknowledge the importance of synthesizing or making new connections at stages in the process of creativity (Osborn, 1957; Parnes, et al., 1977). The second major group, personality and environment, includes those who study creative personalities, social and environmental influences, and affective aspects of personality associated with measurable creative performance. A third group consists of those who consider creativity primarily from its relationship to mental health and full psychological development towards self-actualization (Maslow, 1962). A fourth theoretical perspective stems from the psychoanalytic or psychodynamic theorists who describe creativity as stemming primarily from unconscious or preconscious forces. By his addition of the psychedelic perspective, Gowan (1977) provided a fifth, unique, focus to theories of creativity. He recognized that it is important to practice awareness of the preconscious to tap the normally untouched powerful potential uses of the mind.

The following discussion of orientations to a definition of creativity center on a classification which could be considered to parallel that offered by Gowan (1977) and by Treffinger, Isakson, & Firestien (1983). The parallel classification considers creativity to be more readily studied when these categories are considered: the creative person, the creative process, the creative environment, and the creative product (Arieti, 1976; MacKinnon, 1975).

The Creative Personality

Early research in creativity centered primarily on study of the personality (Arieti, 1976; Barron, 1963) and/or attributes of relatively mature persons who have already demonstrated their creative potential (Heist, 1968; MacKinnon, 1968; Stein, 1974-1975). The historical retrospective approach was described by Taylor (1975) as the study of creativity in creative adults. The method was used in an empirical study of 64 eminent scientists by Roe (1952) and in a study of creative architects (MacKinnon, 1968).

In a review of the literature by Stein (1968), a list of personality characteristics of creative persons was compiled: is an achieving person; has a need for curiosity; has a persistence of motive; is self-assertive and dominant; possesses initiative; is independent, autonomous; open to feelings and emotions; and is not highly critical of himself. MacKinnon (1968) noted that "creatives seem to be continuously self-critical although basically self-accepting, while retaining a sense of destiny, commitment, and involvement in what they are doing" (MacKinnon, 1968, p. 14). The term, "unity of opposites," was used by Torrance (1979, p. 5) as a way to describe an important aspect of creative products and of creative individuals, that is the synthesis or integration of sets of opposites in the personality of highly creative individuals. Highly creative persons, for example, are found to be more masculine and also more feminine, more independent in thought and more open to suggestions from others, more conforming and more non-conforming than persons who are less creative (Torrance, 1979). These characteristics of creative individuals were identified primarily by the historical retrospective method of investigating adult creatives.

The Creative Process

If creativity is defined as a process, then the question of measurement becomes more difficult. For example, selection of the point in the creative process where the researcher shall decide to sample behavior becomes an issue (Getzels, 1975). Dependent upon the particular phase of the process, different techniques might lead to greater complexity of behaviors. How does the researcher view the creative process? The researcher focusing on investigating the creative process could, for example, proceed from a psychoanalytic view (Arieti, 1976) to study the role of incubation of ideas and other unconscious and preconscious influences as inferred events (Barron, 1963; Freud, 1908; Parnes et al., 1977; Vernon, 1970). However the researcher defines the creative process, the task, then, centers on selecting appropriate specific behaviors for observation. The specific behaviors selected will vary depending on the stage of the creative process to be observed, such as phases within the CPS process (Tetenbaum & Houtz, 1978).

Some writers consider problem finding behavior to be central to the creative process. The creative analytic/synthetic process involves disembedding stimuli from their original perceptual field and reorganizing information in a novel manner (Artley, Van Horn, Friedrich & Carroll, 1980). Csikszentmihalyi and Getzels (1971) examined the relationship between the originality of the artist's product and assessment of the artist's "problem finding behavior" before and during the work. They found that artists who showed greater manipulation and exploration of their materials and who deferred closure of form in the initial sketches of their drawings produced paintings which were judged to be more original and to have a higher overall aesthetic value than the art produced by those whose problem finding behavior was more restricted.

Creativity and the Environment

The important role of the environment in the nurturing of creativity has long been recognized. The early environment, which should be considered as a genetic-environmental interaction during infancy, can impact development of creativity through the quantity and quality of stimuli for the child and the attitudes the child develops as the environment is explored (Clark, 1983). Parent child relationships which provide encouragement of independence, of freedom to explore the environment, encouragement of creative talents, and respect for the child appeared to be frequent in the histories of highly creative adults (Christie, 1976; MacKinnon, 1975). Conventional schooling, however, can often create a restrictive environment which can have a negative impact on creativity (Bull, 1978; Clark, 1983; Heist, 1968). The social influences of individual teachers and peers as models can be profound; intimate, encouraging friends can favorably impact motivational orientation (Amabile, 1983; Osborn, 1957). The effect of working with a group to promote creativity has been noted (Amabile, 1983; Arieti, 1976; Hare, 1981; Osborn, 1957; Parnes et al., 1977). Working with peers who might be evaluative or, according to some researchers, simply working in the presence of others, can inhibit creative production (Amabile, 1983; Stein, 1974-1975). Other authors (Bolen & Torrance, 1978; Osborn, 1957; Roweton, 1982) report that interaction with a group enhances creative production. These issues, including effect of competition, are discussed in greater depth under creativity phenomena in groups (see p. 81ff).

When looking at the environment as a variable in creativity research, the question arises: can creativity be increased by training? Or, even, ought creativity training be provided? Torrance's (1972a) assessment of the literature, that it was possible to teach children to think creatively was initially critically questioned (Kogan & Pankove, 1974; Mansfield, Busse, & Krepelka, 1978), but evidence is now available to support the claims (Cohn, 1985; Feldhusen & Clinkenbeard, 1986; Rose & Lin, 1984). A meta-analysis by Rose and Lin (1984) across 46 studies of varied methods of creativity training obtained a relatively low overall effect size which, at first glance, would fit with the skeptical view of creativity training. An in-depth analysis of the verbal scores found moderate improvement of overall verbal scores and a strong impact on verbal originality, thus giving evidence in support of effectiveness of creativity training programs of a verbal nature. The effect of training on figural creativity "needs to be researched more thoroughly" (Rose & Lin, 1984, p. 21). Creative problem solving appears to be the most effective (Mansfield et al., 1978; Rose & Lin, 1984; Torrance, 1972a) of the creativity training techniques and is discussed more fully below.

Creative thinking can be considered to have at least two aspects: as a skill that can be developed through various teaching methodologies, and/or an innate ability that some individuals have in greater abundance than others (Rose & Lin, 1984). In light of this hypothesized dual nature and the relationship of creative expression to mental health (Barron, 1963; Sund & Carin, 1978), it is understandable that the literature abounds with advocates for facilitating the growth of creativity (Davis & Scott, 1971; Osborn, 1957; Parnes et al., 1977; Torrance, 1972a) and particularly for the gifted (Callahan, 1978; Feldhusen & Treffinger, 1985; Gallagher, 1975; Torrance et al., 1980-1981).

Creative Products

In recent years divergent thinking tests have been the most widely used measures in the assessment of creativity (Hocevar, 1981). Hocevar considered the inventory method of listing an individual's creative accomplishments to be the most defensible method to assess creativity. However, young children have seldom had sufficient opportunity to display a number of significant creative accomplishments. This author, then, considers divergent thinking tests to be the most defensible method to assess a young person's ability to tend to perform creatively. Khatena (1982) asserted that it is a myth that we cannot measure creativity especially when we realize that "like intelligence, we are not attempting to measure native creativity" (p. 22). Khatena considered divergent thinking tests to be measures of creative process (1982). However, the tests generate a product, a measure of the person's creative thinking, and, therefore, the process is inferred. This author considers divergent thinking measures to be outcomes of the creative process which are approximate measures of the person's creative potential.

Examination of creative products looks at the finished product of an artist, scientist, or even a young individual where personal commitment was involved in the process (Besemer & O'Quin, 1986; Besemer & Treffinger, 1981). Performance on divergent thinking tasks may or may not reflect much personal comitment. In fact, performance on such tests is susceptible to reflect motivational variables in the testing situation (Elkind, Deblinger, & Adler, 1970).

Measurement of Creativity

A very brief review of the issues, perspectives, and general methodological concerns in the field of measurement of creativity must be addressed prior to any discussion of theoretical or applied issues for specific measures of creativity. The research literature of measurement of creativity centers on four areas of primary concern: 1) issues of construct validation, 2) use of the measurements to assess creative behavior, 3) other aspects of validation of the measure(s), and 4) other methodological issues. These issues are addressed in some depth in Appendix E: Issues in Measurement of Creativity).

Included within the domain of construct validation are the many studies concerned with distinguishing between creativity and intelligence (Carroll & Laming, 1974; Getzels & Jackson, 1962; Wallach & Wing, 1969; Welsh, 1975). The degree to which creativity measurement is related to intelligence is not a concern to the present study. The purpose of the present study in using measures of creativity is to assess the effects of creativity training on intelligent children. Therefore, issues related to the degree of purity of the creativity task in relation to IQ measures need not be addressed further.

Aspects of construct validation relating to the types of creative behavior measured by various instruments and aspects of predictive and concurrent validity of these instruments are relevant. The following discussion is centered on the use of these instruments in research studies of the effect of training, similar to team CPS, on creative behavior.

Multivariate Analyses and Research in Creative and Affective Behaviors

In discussing the areas of research needed to study the nature and measurement of creativity, it was suggested that researchers in creativity become proficient with more complex methodologies (Treffinger & Poggio, 1972). They recommended the use of multivariate statistical techniques to assess a degree of relationship among the various components of the complex behavior of creativity. Multivariate statistics can also control inflation of the alpha level when several univariate analyses of related behaviors are calculated. Some issues related to use of multivariate statistics are discussed in greater depth in Appendix B.

Also suggested were longitudinal studies, "the use of experimental and quasi-experimental designs, including large-scale sampling of populations of interest, well-controlled studies, and replication studies" (Treffinger & Poggio, 1972, p. 259). Now, sixteen years later, there is still a paucity of multivariate studies reported in the creativity literature. Treffinger and Poggio's (1972) recommendations are applicable to the literature of self-concept as another complex behavior.

Testing Creative Thinking

Torrance Tests of Creative Thinking

Content validity. The content validity of incorporating several tasks in the TTCT stems from Torrance's definition of creativity as a complex behavior rather than a pervasive unitary function (Torrance, 1974b). The tasks were selected to sample different creative thinking abilities, as confirmed by factor analytic studies. They are considered to be free of technical or subject matter content. The tasks may be administered to all age levels from kindergarten up. This allows the TTCT to be used with a variety of populations and makes it appropriate for longitudinal research.

Construct validity. Torrance (1974b) considered the intercorrelations between the resultant measures of the tests to be low. The intercorrelation of scores on the dimensions of verbal and figural fluency, flexibility and originality, and of figural elaboration ranged from .02 to .87 for the seven dimensions. The high verbal intercorrelations (.63 to .86) show overlap among the variables, however, thus indicating that a single score could provide satisfactory measurement information (Chase, 1985).

Torrance (1974b) reported studies offering convincing validation of various constructs of a complex definition of creativity. Many of the studies identified children or adults as high or low in creativity and found correlations with a broad range of personality characteristics such as human movement and color responses on the Rorschach and negative correlations on a scale of rigidity and ability to withhold opinion. Correlational studies which did not first preselect Ss on the TTCT, found a relationship between composite scores and an evaluation of originality of imaginative stories. Teacher ratings of children's playfulness and their scores in fluency, flexibility, and originality also were related.

In his review of the TTCT in the Ninth Mental Measurements Yearbook, Chase (1985) failed to mention or acknowledge Torrance's reasoning in presenting a non-factorial, non-simplistic process definition of creativity. He suggested a more meaningful approach to construct validation would be first to sort students on the predicted traits and then to examine their TTCT scores for significant differences.

The new streamlined scores (Torrance & Ball, 1984; in press) change the measures provided by the figural test to an operational definition which reflects a more congruent fit with Torrance's (1974b) definition of creativity (cited in Chapter I, p. 13) than the original scoring (Torrance, 1974b). Torrance (1974b) considered the TTCT to have appropriate content validity to measure various aspects of the definition. He believed this because the activities sampled by the battery of tests are representative of the possible universe of creative behaviors. The problem that he did not discuss was that the scores, fluency, flexibility, elaboration, and originality, reflect a much more limited sample of creative behaviors, that is, divergent thinking, in which originality is defined merely as statistical infrequency.

New scoring criteria and norm tables have been produced for the figural tasks (Torrance & Ball, 1984) and are being developed for the verbal battery (O. E. Ball, personal communications, August, 1985-October, 1987; J. D. Kauffman, February-May, 1987; E. P. Torrance, July-August, 1985; see Appendix A). Standard scores are now available for five aspects of the figural battery: fluency, elaboration, originality, abstractness of titles, and resistance to closure. These new measures provide a greater construct validity for the essential qualities of uniqueness that were hypothesized by Torrance (1979). Criterion-referenced creativity indicators include humor, richness of imagery, and eleven other qualities of a person's ability to express visualizations creatively (Torrance & Ball, 1984).

Streamlined scoring for the verbal battery has been under development for the past several years (personal communications with O. F. Anderhalter, March, 1988; O. E. Ball, August, 1985-October, 1987; J. D. Kauffman, February, 1987, November 1987; E. P. Torrance, August, 1985). When completed, the streamlined scoring will provide standard scores in fluency, flexibility, elaboration, and originality for the verbal form of the TTCT. The use of eight criterion-referenced creative strengths including humor, analogies, future orientation, and richness/colorfulness of imagery, will provide a greater depth of measurement of creativity than the original scoring (Torrance & Ball, in press; Wechsler, 1981). This is true because the streamlined scoring provides a more consistent match with the work of the process theorists and those who stress the role of imagination in creativity, such as Arieti (1976), Gowan (1977), Harrington, Block and Block (1983), Osborn (1957), and Parnes et al. (1977). The creative strengths have been empirically demonstrated in a factor analytic study with the figural tests (Mourad, 1976). Further discussion of prior validity evidence for the streamlined figural scoring procedures is presented in Chapter III and Appendix E.

Predictive and concurrent validity. A pattern of lower reliability and of concurrent validity scores for the figural tests (standard scoring), when compared to the verbal tests, emerged in this reviewer's analysis of the literature. This problem might be corrected by use of the figural streamlined scoring system recommended by Torrance and Ball (1984). For a fuller discussion of issues concerning the streamlined scoring, see Appendix E.

Evidence that TTCT scores predict future creative behaviors is presented in several studies (Howieson, 1981; Torrance, 1972b, 1974b; Torrance & Ball, 1984). These studies are open to the basic criticisms of the creativity research cited in the discussion below; the studies present several univariate correlational analyses of the data, and the scores are primarily measures of divergent thinking. Multiple univariate analyses are subject to inflated significance levels of chance effects, and do not have the power to establish the relative strength of the dependent variables as can be done with a multiple regression analysis. On the other hand, the now outdated scoring system of the figural TTCT was limited as a measure of divergent thinking. The standard scoring of the TTCT missed other qualities of creative behavior, which could deflate the predictive relationship between the test and the criterion variables.

The measurement of creativity in longitudinal studies by Kogan and Pankove (1974) differs from that cited by Torrance (1972b) and from Howieson (1981). Kogan and Pankove utilized the Wallach-Kogan divergent thinking measures rather than the TTCT. Kogan and Pankove described the Wallach-Kogan measures to be independent of IQ. In their 1974 study, the fluency scores were used as one of the predictor variables in a multiple regression analysis of overall nonacademic activities and achievements. Kogan and Pankove found fluency scores at tenth grade to be of limited predictive value, and at fifth grade to be of no demonstrated predictive value for nonacademic activities and achievements of students in their senior year. The present author sees two serious problems which were not discussed by Kogan and Pankove. The authors excluded uniqueness scores as a predictor variable because they asserted that fluency and uniqueness were highly correlated variables with a median correlation of .72 (Kogan & Pankove, 1974). This author thinks that uniqueness might have been a logical predictor variable of the criterion of extracurricular activities and should, therefore, have been included. A more serious difficulty with their study is that the criterion variables show questionable stability in their analysis. The tenth grade level of student activity was included in the twelfth grade cumulative autobiographical inventory of student activities and achievements and, therefore, the predictive power of the tenth grade activities is not independent of the dependent measure, cumulative activities. Their conclusions were that the fifth grade measures of ideational fluency offered little of predictive value of nonacademic achievement in the high school years. However, the "lack of predictive validity does not necessarily imply the absence of concurrent or construct validity . . . There is good reason to believe that fifth-grade children with high scores on ideational productivity differ from those with low scores on a variety of cognitive and personality measures" (Kogan & Pankove, 1974, p. 808).

Torrance and Safter (1986) indicated that studies of long term predictive validity of scores on the TTCT showed substantial relationships to adult creative achievement. However, children must practice the acquired creative behaviors or the "gains in creative functioning are unlikely to be retained" (Torrance & Safter, 1986, p. 7). In their study, they found that mean creativity test scores of fluency, flexibility, and originality increased from 1976 to 1982 when new norms were computed for the TTCT verbal scores. They suggested that the increase in creativity scores reflected an increased focus on teaching of creative thinking skills incorporated in textbooks, teaching methods, statements of curricular objectives, and instructional materials (Torrance & Safter, 1986).

Similes Test

Construct validation to establish the Similes test as a measure of verbal rather than figural creativity shows correlations, $r=.32-.58$, of Similes scores to verbal fluency scores from three verbal TTCT subscales, and low correlations, $r=.02-.14$, with figural fluency scores from two TTCT figural tests (Schaefer, 1971), and $r=.18$ with a multiple choice test of preferences for similes (Pearson & Maddi, 1966). Torrance and Ball (1984), however, reported significant correlations of Similes to figural fluency, .42; to figural originality, .40; and to abstractness of titles, .35. Correlations with verbal and nonverbal measures of achievement and IQ in the range from .32 to .41 were reported by Schaefer (1971). Schaefer cited other validity studies which show that Similes scores discriminated significantly between elementary children identified by their teachers as performing creative work and those who did not. Schaefer reported that for a sample of college women, a correlation of .31 was reported with the score on a biographical inventory assesing creative skill in writing.

Use of Similes is an appropriate measure of creativity because the concepts of relevance or essence as addressed in its scoring of originality. It provides a needed dimension to evaluate the validity of the new streamlined scoring variables of the TTCT figural creativity tests. These measures are appropriate to assess the type of creativity needed for effective CPS such as in OM where teams must find creative ways to solve a real problem.

Creative Problem Solving

The OM program provides training and competition experiences in team creativity using CPS procedures. As seen below, the effectiveness of CPS as a training procedure has been documented in the literature. The next sections of this review focus on CPS as a process, OM and Future Problem Solving as CPS programs, a summary of research to identify personality characteristics and CPS training, and on team CPS and gifted children.

The Process of Creative Problem Solving

Early authors in creativity and gifted education and in problem solving sought to identify the creative process involved in the solution of complex problems. In Wallas' (1926) paradigm of the creative process, the role of the preconscious was considered to be important. He suggested four phases: preparation (including gathering information), incubation, inspiration, and evaluation (elaborating and assessing the solution). Osborn (1957) considered the creative process for solving problems to usually consist of seven phases: orientation (problem awareness); preparation; analysis; ideation (the production of alternative ideas); incubation; synthesis and evaluation. The steps in the phases can, and indeed must, vary with different problems.

Recently the term creative problem solving (CPS) has been more consistently applied to the situations in which the complex problem solving process is structured to require creative thinking at given stages. Maker (1982) differentiated Parnes' CPS model from usual problem solving methods in that it emphasizes production of diverse alternative ideas before selecting a solution or developing a plan to implement and find acceptance for a solution. The Parnes method of CPS (Parnes et al., 1977) was developed to provide a comprehensive, theoretically sound, and effective process to stimulate the use of imagination in real situations.

The manner of statement for the problem is an essential determinant of whether the problem will be conducive to CPS. It is essential that the problem be stated in open-ended and ambiguous terms to reduce functional fixedness, preconceived ideas about the problem (Gourley & Micklus, 1978; Micklus & Micklus, 1987). Parnes et al. (1977) referred to the problem situation as "the mess," (p. 88) and Crabbe (1985) described the problem of Future Problem Solving as the "fuzzy situation" (p. 15).

In an analysis of problem finding behaviors, Getzels (1975) distinguished between problems in which the solution is known to the question framer and those in which the solution is not known. For a problem situation to be considered for CPS, solutions must not be known and the problem cannot be solved by only one correct answer. Problems with a single correct answer are considered convergent rather than divergent problems (Guilford, 1967; Meeker, 1969).

Parnes et al. (1977) credit Osborn (1957) as the first to introduce the concept of deferred judgment. To Parnes et al., the principle of deferred judgment is essential to the CPS process. By withholding an evaluation of ideas during the idea production phase, the person's imagination is freed to work on the problem, and if this is done in a nonevaluative atmosphere, there is usually a greater willingness to take risks. Persons learning the CPS process by Parnes' methods receive instruction in awareness of and use of the incubation phase of the problem solving process as an extension of the principle of deferred judgment (1977). The number of steps within the CPS process can vary: five are seen by Parnes et al. (1977), six by Isaksen and Treffinger (1984), or eleven by Bull (Bull & Fishkin, 1987). Two CPS models specifically adapted for children both specify six stages, but vary in conceptualization of the process at some of the stages (Eberle & Stanish, 1980; Torrance et al., 1980-1981). There appears to be little consensus on the consistency of the steps in the CPS process. However, it could reasonably be inferred that theorists of the CPS process agree that it is a complex, multi-faceted process.

The following review of the CPS literature is limited to those studies that specify the CPS process as problem solving situations meeting the following descriptions: 1) The problem, as stated, must be phrased in such a way that more than one solution would correctly solve the difficulty (Bull & Fishkin, 1987; Getzels, 1975; Micklus & Micklus, 1987). 2) The process stresses brainstorming or divergent thinking (Osborn, 1957; Torrance et al., 1980-1981), and the principle of deferred judgment is used during the divergent phases which alternate with the evaluative process phase at the various steps (Firestien & Treffinger, 1983; Parnes et al., 1977).

Effectiveness of Creative Problem Solving

Osborn (1957) noted that the techniques of deferred judgment, brainstorming, checklisting, and setting aside a time and place for creative thinking helped to produce valuable ideas. He also noted that, with a focus on student participation and practice of the process, people became more open to ideas and seemed to appreciate their own deep well of knowledge, both public and private (Parnes et al., 1977).

Extensive research in the effectiveness of CPS has been conducted by Parnes and his associates at the Interdisciplinary Center for Creative Studies at State University of New York (SUNY) at Buffalo (Parnes et al., 1977; Rose & Lin, 1984). The early research from 1949 to 1956 under Osborn's direction focused primarily on the pilot study of a one-semester creativity training program. The program evolved as Parnes' conceptualization of CPS as a training process developed. Significantly more ideas and good quality ideas were produced by subjects trained in the principles of deferred judgment than by those who did not take the CPS course (Parnes, 1961). Early findings indicated that students maintained benefits of the semester course in CPS even eight months later by scoring significantly higher on tests of creative thinking abilities than untrained subjects (Parnes & Meadow, 1960). Meadow and Parnes (1959) studied personality characteristics as well as creativity measures and found scores to increase on one of the four measures of the California Psychological Inventory, the scale of dominance. It was found that a greater proportion of good ideas were produced among the ideas later in problem solving session (Parnes, 1961).

In a critical review of the effectiveness of the Parnes' CPS program as well as four other methods of creativity training, Mansfield et al. (1978) located methodological deficiencies in several of the studies. Mansfield et al. noted that there were significantly greater gains in the instructed students than the controls on two measures of quantity of ideas, three of the five measures of quality of ideas, and on a dominance scale. However, the validity of the Meadow and Parnes (1959) findings were open to question because of a failure to assign subjects to the two groups randomly. Mansfield et al. considered that the Meadow and Parnes (1959) study, as well as later studies, permitted the instructed students to receive practice in divergent thinking tasks similar to the criterion tasks.

Mansfield et al. critically reviewed four other studies by Parnes and his associates. Of the studies they reviewed, they considered Reese and Parnes' (1970) study to be one that was well-designed. High school senior volunteers participated in a self-instructional program, instructor-led program, or a control group. The instruction occurred two times a week for 26 sessions during the students' study period. The results were considered by Mansfield et al. to be consistent and impressive. They concluded that six evaluation studies of courses using the Parnes material show evidence of effectiveness and greater success than evaluations of either the Purdue Creative Thinking Program (Feldhusen, Speedie & Treffinger, 1971) or the Productive Thinking Program (Treffinger & Ripple, 1969). Mansfield et al. commented that the diversity of technique in the Parnes program contributed to success of the program but made it impossible to determine which specific techniques were effective. Three additional studies by Parnes' associates in 1959 and 1965 were cited which showed effectiveness of the brainstorming technique alone when used in only a single session. Mansfield et al. speculated that the greater success of the Parnes CPS program over other creativity training programs may be due in part to the breadth of training techniques, the incorporation of brainstorming practice, and the exclusive use of the CPS program with persons of high school age or older. In Parnes' work, the control subjects, as well as those who were instructed, were selected from volunteers who registered for the course. The effectiveness of the CPS program, then, may be enhanced by prior motivational factors in the subjects.

Mansfield et al. (1978) noted a number of methodological problems common to evaluation studies of creativity training. In several studies, the sample size was small and consisted of intact classes assigned to each condition. In those smaller studies, extent or manner of teacher involvement in the training is confounded with the training program per se. Other common problems included the use of several univariate analyses rather than a single multivariate analysis; failure to control for Hawthorne effects (effects on motivation of the subjects due to knowledge of participating in a study); failure to assign subjects or classes randomly to treatment conditions; and the use of individual scores rather than classroom means as the unit of analysis.

Mansfield et al. viewed Torrance's (1972a) conclusion that creativity can be trained as premature "since conceptual and methodological problems pervade most evaluation studies" (p. 517). They viewed the chief conceptual problem to be equating performance on tests of divergent thinking with creative behavior. This issue was addressed under ways of defining and measuring creativity. The problem of similarity of the training procedures to criterion evaluation measures may perhaps be less of a problem with CPS training methods than other procedures because of the variety of techniques used in CPS.

Cohn (1985) conducted a research synthesis of 106 published studies and dissertations on creativity training. By systematic analysis of the coding of the data from the studies, Cohn was able to calculate effect size statistics for fluency and originality of responses. Her findings suggested that creativity training can increase creative responses, however, similar changes can be produced by varying warm-up procedures or other motivational conditions. Cohn's results also indicated that creativity training effects are less pronounced when the criterion tasks used to assess the training are dissimilar to those employed during the training.

Torrance (1972a) evaluated 22 studies which used the Osborn-Parnes CPS program and an additional 120 studies using other creativity training methods for elementary and high school students. All of the nine methods of training had more than a 60 percent success rate, and the CPS program method showed a greater percentage (91 percent) of success than any of the other methods he studied. Success was defined as a numerical rating representing the proportion of objectives of the experiment which reached significance at the .05 level of confidence. Torrance summarized his findings by saying:

> . . . it does indeed seem possible to teach children to think creatively. The most successful approaches seem to be those that involve both cognitive and emotional functioning, provide adequate structure and motivation, and give opportunities for involvement, practice, and interaction with teachers and other children. Motivating and facilitating conditions certainly make a difference in creative functioning but differences seem to be greatest and most predictable when deliberate teaching is involved (Torrance, 1972a, pp. 132-133).

The success of CPS methods found by Mansfield et al. (1978) and by Torrance (1972a) was corroborated by the recent meta-analysis of long-term creativity training programs conducted by Rose and Lin (1984). They established a relatively homogeneous research population of long-term creativity training programs by limiting their sample to studies which evaluated a series of lessons or on-going treatments and only to studies which used the Torrance Tests of Creative Thinking (TTCT) or its modified form as the criterion instrument. They considered the TTCT to be an acceptable measuring instrument whereas others were considered to be "questionable" (Rose & Lin, 1984, p. 16). The use of the TTCT helped to establish a consistent operational defintion of creativity across the 46 studies which were analyzed.

The effect of creativity training programs on TTCT posttest scores across all 46 of the programs would be classed as a small effect size (ES = .468) according to Cohen's (1977) standards. The majority of the programs had a high reliance on verbal activities in the training and it was not surprising that treatment effects for verbal TTCT scores (ES = .596) was moderate. The training program found to have the most consistent impact on TTCT scores was that of Osborn-Parnes CPS. The overall positive effect size for CPS (ES = .629) was moderate with training explaining over 40 percent of the variance in subjects' scores. In CPS, as well as most of the other types of training, the greater effects were noted on originality and fluency scores, both verbal and figural. Verbal flexibility scores were also strongly affected by the CPS treatment. In general, verbal creativity was more strongly affected by these programs than was figural creativity. Rose and Lin suggested that it is possible that the figural form of the TTCT could be measuring aspects of the individual's creative ability that is not readily affected by training.

Rose and Lin (1984) were discouraged by the paucity of systematic research in program evaluation studies in creativity. Evaluations of the Osborn-Parnes CPS program were the most systematic and voluminous found in the literature, yet because of the limited number of studies in the meta-analysis were not "sufficient to warrant an unreserved opinion of the program's effectiveness" (Rose & Lin, 1984, p. 21). Rose and Lin's in-depth examination of the differences in programs, of specific TTCT score categories, and of the impact of CPS on verbal creativity, and the findings of other researchers (Parnes & Brunelle, 1967; Torrance, 1972a) provide "strong evidence to support the effectiveness" of the CPS program (p. 21). However, an excessively large effect size was found for verbal originality and should be viewed with some caution. The data in the meta-analysis are conducive to inferences that programs which provide "more varied and flexible experiences" (p. 22) essential to creativity are among those with a greater treatment effect.

In a recent review of research related to creativity instructional materials, Feldhusen and Clinkenbeard (1986) described the major creativity training programs and discussed research concerning their effectiveness. They focused primarily on the Purdue Creative Thinking Program, Productive Thinking Program, CPS when used with adults, and some additional creativity material including New Directions in Creativity (Renzulli, 1973). They concluded that the results clearly indicate that it is possible to "effect significant gains in students' creative thinking and problem solving abilities particularly as measured by divergent thinking tests" (Feldhusen & Clinkenbeard, 1986, p. 177). However, they considered tests of divergent thinking which only yield scores of fluency, flexibility, originality, and elaboration to be limited in defining creativity. Moreover, when only such tests are used as an evaluation of programs using divergent thinking materials, other complex aspects of productivity and related affect are not assessed. Based on a study by Harrington et al. (1983), Feldhusen and Clinkenbeard suggested that "one solution to the validity problem may be to use a 'high quality' response score for divergent thinking tests" (1986, p. 177). This is an argument in favor of the now available creative strengths scores of the streamlined figural scoring for the TTCT (Torrance & Ball, 1984). They found some support also that creativity training causes positive effects on related attitudinal measures (Reese, Parnes, Treffinger & Kaltsounis, 1976; Shively, Feldhusen & Treffinger, 1972; Treffinger & Ripple, 1969). Feldhusen and Clinkenbeard suggested use of a wider variety of criteria including affective or personality dimensions, relationships between creativity and independent learning, and "attention to the persistence necessary to develop a creative product" (1986, p. 178).

The preceding review shows strong evidence for the effectiveness of CPS as a training procedure for increasing creativity scores. Evidence for the effectiveness of CPS on affective measures is less readily available (Meadow & Parnes, 1959; Shivley et al., 1972). Although educators and psychologists recommend that the CPS process should be learned and practiced as a method to promote mental health (Parnes, 1975; Parnes et al., 1977; Sund & Carin, 1978) and leadership abilities (Osborn, 1957; Parker, 1983), there is still very little research evidence to support such claims.

Affective traits and their relationship in gifted children who exhibit good or poor skills at CPS were studied by Sherman (1977), Houtz, Rosenfield and Tetenbaum (1978), and Tetenbaum and Houtz (1978). These studies, which did not examine the effect of training in CPS, are discussed below under studies of personality and CPS, and locus of control expectancies and the gifted.

Team Creative Problem Solving for Children

Odyssey of the Mind

Odyssey of the Mind provides CPS training for students, ages kindergarten through high school (Gourley & Micklus, 1978). The problems in OM are phrased in an open-ended manner to reduce functional fixedness (Micklus & Micklus, 1987) and redefinition is encouraged. Brainstorming techniques, the principle of deferred judgment, and recognition of the value of the incubation process are utilized (Bull & Fishkin, 1987; Micklus & Micklus, 1987; Moyers, 1981). OM training materials for coaches consistently use a variety of CPS techniques. The training of coaches in Oklahoma OM (OK-OM) program specifically uses a CPS model developed by Bull (Bull & Fishkin, 1987) as a general problem solving model and lists ways to adapt the model with students.

A personal communication with the co-founder of the OM program affirmed that as of November, 1987 (Micklus, personal communication, November 11, 1987, see Appendix A) no research studies measuring changes in behavior on OM team members have been completed. Four studies utilizing a variety of questionnaires as instruments have evaluated data from students, teachers, coaches, and parents (Christy, personal communication, July 14, 1988; Goff, 1987; Harrington, 1984; Miller, 1983). A fifth study (Cohen, 1987) provided a case approach with behavioral data on team members and coaches of five teams. However, none of these studies assessed changes in the subjects.

Harrington (1984) conducted a master's thesis research survey of 250 students in grade three through twelve who had participated in OM. Students responded to a survey of how the program affected their problem solving skills. More than 50 percent of the students felt that OM helped them to increase in risk-taking, close to 50 percent felt they increased in flexibility. A large percentage (91%) indicated they learned better problem solving skills and definitely wished to try the program again (80%).

Teachers and coaches rated the program very high in the areas of originality, fluency, taking risks, learning from mistakes, working with others, and helping students to solve problems effectively.

> Both students and teachers felt the program was challenging and required hard work and persistence (p. 16). Overall, the problems encountered were far outweighed by the good things that the program seemed to accomplish . . . Students were able to work together, they shared ideas, made new friends and they learned some valuable skills while doing it in an atmosphere that was fun and enjoyable (Harrington, 1984, p. 64).

Harrington limited her data observations to reports by others and the team members. She did not use direct measures of student behavior such as creativity or problem solving tasks.

Olympics of the Mind (now Odyssey of the Mind) was a component of gifted programming in an elementary gifted program evaluated by Miller (1983). Miller's results showed that the students appeared to be experiencing success in VITAL (the gifted program), especially in the OM program. Miller's (1983) study, like Harrington's (1984), used descriptive methodology. Miller's instruments, however, were somewhat more varied. She used results of surveys; questionnaires from teachers who taught the students in cluster groups in their regular classes, parents, and students in the gifted program; and interviews with students who recently graduated from the program. Additionally, interviews were conducted by an outside evaluator with the gifted specialist teachers, administrators, and representative classroom teachers, parents, and students. Based on her findings, Miller recommended the following:

> The VITAL staff should explore ways of incorporating elements of the successful Olympics of the Mind program, e.g., competition, group problem solving, creativity, into other VITAL activities. Aspects of the Olympics of the Mind program might also be successful in use with non-VITAL students. The VITAL staff should continue the challenging activities which are being used in the VITAL program. Olympics of the Mind appears to be particularly challenging (Miller, 1983, p. 95, 97).

In a study designed to reveal perceived definitions of creativity among people associated with a gifted program, Goff (1987) compared groups who were involved in OM to those who were not. Because of the small number of those involved in OM (10% of the entire sample), the responses of the three adults (two judges and one parent) were added to those of the 22 children who were in OM. A non-OM control group with comparable number of students, teachers, and parents was randomly selected from the non-OM group. Of the five factors identified by a factor analysis of the 14-item Likert-type creativity survey, only one had adequate reliability. The factor considering creativity to be a teachable characteristic had Cronbach's $\alpha = .68$. The finding of low reliability is not surprising considering that there were so few items within each factor.

The results failed to find a difference between the OM groups or between males and females in perceived definitions of creativity. It is unknown whether there were differences in perceived definitions of creativity between the adults and children in the study. With the possibility of differences in perceptions existing between the adults and children and low reliability of the five factors used as dependent variables, it is very possible that real differences which could exist would not emerge. The study by Goff (1987), with exclusive use of a single self-report measure, is subject to problems of validity associated with the Harrington (1984) study (Isaac & Michael, 1981).

Perceptions toward creativity as a socially acceptable and a desirable trait for children are under investigation by N. R. Christy (personal communication, July 14, 1988, see Appendix A). She is surveying adults in Oklahoma, parents, teachers and/or coaches from three sources: from winning teams at the State OK-OM competition, from teams who performed at OM competition but did not win at the State OK-OM finals, and from adults associated with children who never participated in OM. Christy surveyed the three groups of adults with three kinds of self-report questions: attitudes toward nurturing of creativity in schools and at home, attitudes toward creative behaviors in one's own children/students, and perception of personal creativity utilizing a creative self-concept scale, Ideas by Fishkin (1987a, 1987b). In an assessment of some preliminary findings, Christy indicated that there appear to be differences in perceptions toward creativity between the coaches and the parents of winning teams. That is, many of the parents appear to have a more restrictive or controlling view than the coaches of the acceptability of creative behaviors. The response rate from adults associated with teams who competed but did not win at state was much lower than that from the winning teams or from non-OM groups. A tenuous interpretation of the smaller sample (at present only 20) of responses from non-winning OM groups indicates that this group may have different attitudes toward creativity than the other groups (N. R. Christy, personal communication, July 14, 1988).

A descriptive study in a case study format was conducted by Cohen (1987). Five OM teams and their coaches were studied by means of interviews, self-report surveys, school records, and instruments to measure creativity, the verbal and figural batteries of the TTCT and affect, the Piers-Harris Children's Self-Concept Scale, for the children and the Adjective Check List for the coaches. No inferential statistics were used to make comparisons. However, examination of the descriptive data permitted Cohen to make observations about some of the likely similarities and differences among the five teams.

Three of the five teams were winning teams at their state competitions, two of which won first place and advanced to world competition. All teams showed higher scores on verbal creativity than on figural. Unlike the other three teams, the two championship teams showed TTCT verbal creativity scores greater than the 90th percentile for most of their members. It is possible that the TTCT scores should be considered as posttest scores because the measures were administered at least one to two months after the teams had begun their work. Therefore, the higher scores in verbal creativity may be reflecting the better training experienced by those team members. Cohen noted that the championship teams "spent more hours per week working on OM than the non-champion teams" (Cohen, 1987, p. 228), thus implying a likely relationship between winning and effort, and possibly between higher scores in tested creativity and effort. Cohen also observed tentative relationships between higher creativity scores, including the specific categories of creative strengths and subscales of the TTCT, between the coaches and their team members. Most of the team members enjoyed their OM experience and wished to participate again. The successful coaches were those who scored high on the personality characteristics of "Achievement, Self-Confidence, and Creative Personality {or} . . . Personal Adjustment" (Cohen, p. 225), which are among the personality attributes consistently noted in highly creative adults (MacKinnon, 1968; Schubert & Biondi, 1977; Stein, 1968). However, it should be noted that these observed relationships are based on only a very small number of cases in this case study and, hence, should be considered as a very tentative base for further study.

To date, then, there have not yet been any studies to compare the effect of OM on children's creativity or on any aspect of their affect. The findings of the studies by Cohen (1987), Harrington (1984), and Miller (1983), indicate that OM is perceived to be a successful, motivating program in which the children feel challenged, worked together as a team, and in which many felt they had worked hard.

Future Problem Solving

In work on Future Problem Solving teams, students learn and practice the CPS technique and process in a structured manner. Six steps are used: brainstorming problems, analyzing and synthesizing a problem statement, brainstorming solutions, determining criteria for evaluation, evaluating the better solution according to the criteria, and refining the best solution (Crabbe, 1985; Torrance et al., 1980-1981). Primary (second and third) grade students do not practice all levels (see Chapter III and Appendix N). Although the Future Problem Solving program has been asserted to be beneficial to children in student testimonial (Crabbe, 1982) and teacher observations support the program (Hoomes, 1984), according to an ERIC search (April, 1988) only three research studies have as yet been completed.

Tallent (1985) compared elementary gifted children who were experienced in the FPS program technique and those who had not experienced the training. The children were given an FPS problem packet to work on an individual basis. The packets were evaluated by trained FPS evaluators using FPS program evaluation criteria. The results showed strong support favoring the group trained in FPS programs (Tallent, 1985).

Torrance and Mourad (1978) conducted a descriptive pilot study of characteristics of 1,729 students who participated in the Future Problem Solving program as an aspect of their gifted programs. The study was further described in an article by Torrance (1978). Students in grades 3-12 responded to two instruments, a self-rating Likert-type scale, "Abbreviated Self Directed Learning

Readiness Scale," and an abbreviated form of "Thinking Creatively About the Future." The 10 items of the self-rating scale were selected by Torrance primarily from the Creative Learning factor items of Guglielmino's full scale for self-directed learning (Guglielmino, 1977, cited in Torrance & Mourad, 1978). There was a strong correlation between the scores on self-directed learning readiness and originality for all grade levels and the readiness score and fluency for eighth grade and grades 11 and 12. There were indications of developmental growth in the increasing scores of most of the items of the self-directed readiness scale. The results suggested that students in the gifted programs "become increasingly confident in their ability to work on their own but not as a member of a team in solving problems" (Torrance & Mourad, 1978, p. 185). It was noted that a limitation of the study was the lack of a control so that the findings merely provided some useful baseline measures of students in gifted programs who had some experience in the Future Problem Solving program.

It is possible that generalizability of the findings of Torrance and Mourad (1978) and of Tallent (1985) may be affected by a problem common to many studies of creativity training, those where the training procedures are very similar to the criterion variable. It is difficult with the known information to determine how much of the effects were due to a generalized boost in creativity or merely due to practice and familiarity with the techniques of FPS.

Hendrickson (1985/1986) conducted a study of the effects on gifted children who participated in three varieties of CPS programs, including Future Problem Solving. The 58 students were randomly placed in three treatment conditions with 12 teachers. Treatment consisted of: a) practice with two or three rounds of the Future Problem Solving program as within state Bowl competition, b) a combined curriculum consisting of Future Problem Solving Bowl competition combined with Renzulli's Type III real problem, product oriented CPS in preparation for competition at a district-wide student creative products fair, and c) CPS alone in preparation for the creative products fair (Hendrickson, 1985/1986).

The teachers who coached their teams for Future Problem Solving competition (groups a and b) completed 20 hours of training in three day-long workshops. Those teachers who worked with the students in groups b and c attended a 10-12 hour inservice on CPS. The in-class training sessions for children in groups b and c included preparation for the creative skills fair as the primary focus. However, the teachers in the CPS training groups "encouraged students to use the process . . . to insure generalization" (Hendrickson, 1985/1986, p. 72).

The dependent variables were products and performances at the district-wide skills fair scored for resolution and elaboration or synthesis using definitions of product judging from Besemer and Treffinger (1981). The hypothesis that treatment (b) would be most effective followed by treatment (c), and finally treatment (a) was supported in that students in the combined curriculum created more products than students in either of the other groups. Student and teacher mastery of problem solving skills and student mental ability were significant covariates in the analysis of covariance of product elaboration/ synthesis. Younger students contributed more products than older students.

Hendrickson's (1985/1986) study made several important and unique contributions to the literature on team CPS and gifted children. It is the only evidence located by this reviewer to evaluate the effects of a <u>team</u> CPS rather than individual CPS programs and one that is used in a sample of gifted children. By using the criterion of products scored for quality, Hendrickson's study appears to have bypassed the frequent criticism of construct validity problems in many of the preceding studies, which used tests of divergent thinking. His product rating system shares consistency with those authors who suggest looking for the essence in creativity (Besemer & Treffinger, 1981; Harrington et al., 1983; Torrance, 1979; Torrance & Ball, 1984).

It is interesting to note that Hendrickson's findings in his study of his coaches are consistent with the tentative findings of Cohen (1987). That is, a relationship was found between characteristics of students and teachers or coaches. However, Hendrickson's relationships were specific to correlations in mastery scores on a CPS success test administered to teachers and students. Teacher experience and/or performance on the CPS test also "contributed significantly to explaining variance for product total, novelty, and elaboration/ synthesis" (Hendrickson, 1985/1986, p. 111) as analyzed by multiple regression procedures.

Hendrickson's (1985/1986) findings are consistent with the literature supporting the effectiveness of CPS as a creativity training program. The finding that the most complete of the experiences, combining the team CPS of Future Problem Solving with thorough training in individual CPS processes, geared toward solution of real problems was the most effective fit with findings by Rose and Lin (1984) and Torrance (1972a). It is not clear whether the finding that participation in the Future Problem Solving Bowl competition was the least effective of the three CPS training programs might speak to issues of evaluative effects of competition on creativity (see discussion below), or to the potency of CPS training for real problems that is deliberately focused on generalization of the CPS skills as an effective complex training program.

Team Creative Problem Solving and Gifted Children

The focus of this study is to examine the effects of team CPS in a sample of gifted children. It is necessary, then, to establish the appropriateness of team CPS to the education of gifted children.

Participation in CPS work is a curriculum modification which can be shown to be directly related to many of the cognitive, affective, intuitive, and societal needs of the gifted. In an extension of the work of Seagoe (1975), Clark (1983) related the needs of gifted children to the characteristics which differentiate them from other learners. Of the curriculum needs described by Clark which pertain to the specific characteristics of gifted children, team CPS offers differentiated education to appropriately address the following:

1. Because of their high ability to synthesize and systematize, gifted children need increased time to permit incubation of ideas.

2. Because of their high ability to generate original ideas and solutions, gifted children need to develop skills in problem solving, creative thinking, and helping in the solution of meaningful problems.

3. Because they have an "advanced ability to use and form conceptual frameworks" (Clark, p. 93), gifted children need to design and use systems for gathering information and solving problems. They also need to become more tolerant of ambiguous phenomena.

4. Because they tend to be critical evaluators of themselves and others, gifted students need to develop skills in evaluating data, deferring judgment when appropriate, making decisions, and to experience group interaction with individuals who possess varied ways of seeing and solving problems.

5. Because the gifted tend to be acutely aware of themselves to the extent of feeling different from others, gifted students need cooperative learning opportunities in which they can assertively express their needs, feelings, and ideas, in order to share themselves with others and to clarify and affirm their ideas and feelings (Clark, 1983).

The team approach to CPS in OM and in Future Problem Solving creates an opportunity for highly interactive work in which the children become involved with each other in real and/or future-oriented problems. The children exercise skill in discussion, brainstorming, decision-making, and leadership. It is helpful to them to learn to defer judgment and thereby separate the evaluative stages of problem solving from the idea production stage. Because creativity is a necessary skill in all areas of endeavor, gifted students need the techniques and practice involved in CPS to develop their creative potential. Furthermore, the team process is especially helpful to their societal needs for "encounters with social problems, awareness of the complexity of problems facing society, conceptual frameworks for problem-solving procedures; meaningful involvement in real problems; and an understanding of various leadership steps and practice in leadership skills" (Clark, 1983, pp. 95-97).

Because team CPS so appropriately addresses the needs of gifted children as an appropriate differentiated modification of curriculum, many authors have been calling for it to be used in gifted programs (Crabbe, 1982; Houtz, Rosenfield & Tetenbaum, 1978; Parker, 1983; Torrance, 1985). However, with the exception of the study by Hendrickson (1985/1986), the effects of training in team CPS with gifted chidren were unknown because there had been no prior research comparisons. The majority of the studies on CPS which yielded favorable posttest scores compared to pretest scores were conducted with adult subjects or with children who were not previously selected for a gifted program (Maker, 1982; Torrance, 1972a).

Studies of Personality Characteristics in Relation to Creative Problem Solving

Several experimental studies of CPS in gifted children were conducted by Fordham University scholars (Houtz & Speedie, 1978; Sherman, 1977; Tetenbaum & Houtz, 1978) which examined the relationships among various personality characteristics associated with performance of CPS tasks. In these studies, the children completed an affective self-report instrument(s) and individually responded to a variety of problem solving tasks. The tasks, involving divergent and problem solving behavior, were verbal maze problems and written simulation exercises (Houtz & Speedie, 1978). Sherman (1977) used problem situations designed to evoke CPS skills at three different stages of the CPS process. Student scores on the various CPS tasks were then related to student characteristics to determine the relationship of the characteristic under study to skills needed in that step of the CPS process. Of these studies of creativity and problem solving behavior in children, only those which used the population at Hunter College Elementary School for Gifted Children and which investigated affective traits are reviewed (Houtz, Rosenfield & Tetenbaum, 1978; Sherman, 1977; Tetenbaum & Houtz, 1978). Other studies used intact classes of normal students in investigations assessing a variety of tasks in divergent and problem solving behavior (Speedie, Treffinger & Houtz, 1976) or establishing relationships among measures of creative thinking, intelligence, and evaluation skills using the Purdue Elementary Problem Solving Inventory (Houtz, Montgomery, Kirkpatrick & Feldhusen, 1979).

The studies by Tetenbaum and Houtz (1978) and Houtz et al. (1978) utilized nine problem solving measures and three measures of affective characteristics of the students from intact classes at an elementary school for gifted children of varied socioeconomic and racial backgrounds. The affective instruments were the Bialer-Cromwell Locus of Control Scale, a shortened form of the Coopersmith Self-Esteem Inventory, and the Rydell-Rosen AT20 which assessed tolerance of ambiguity. The CPS measures were six measures focusing on divergent aspects, hypothesis generation, and three measures of the convergent, hypothesis-testing-evaluation tasks which involved rearrangements of various problem elements and statement of a final goal.

Tetenbaum and Houtz (1978) performed a factor analysis of the nine CPS tasks which yielded two factors: fluency and rearrangement (use of evaluative and synthesizing skills). An ANOVA by grade and sex showed no sex differences on the affective measures and no grade differences on locus of control or self-esteem. An ANOVA of the CPS factor scores showed that girls were more fluent than boys and boys did better on tasks requiring rearrangement. The authors stated that sixth graders were significantly different from fourth graders in rearrangement but did not specify the direction of the difference.

One focus of the Houtz et al. (1978) study was the investigation of developmental growth patterns in creative thinking and problem solving skills and of the relation of achievement and affective variables to those skills. Each child's score was classified as being in the upper, middle, or lower third of performance on eight creative and seven problem solving tasks. A pattern of growth was seen for all of the problem solving tasks except the verbal maze problem. A pattern of continuous growth from second to sixth grade was observed for the problem solving or convergent skills to a greater extent than for the creative or divergent tasks. Creativity growth was most pronounced between Grades 3 and 4, with no differences between Grades 2 and 3, or Grades 4 and 6. A factor analysis found two separate factors, achievement and fluency of ideas, thus lending support to other researchers' findings that these skills can be separated conceptually (Guilford, 1967; Houtz & Speedie, 1978). When differences in achievement among the children were controlled by means of statistical analysis, internal locus of control and high self-esteem were more significantly related to the creativity and problem solving measures than was tolerance for ambiguity. The evidence of the relationship of affective measures to problem solving performance

> . . . suggests that factors other than achievement in school subjects may be important to the creative and problem solving process and perhaps should be attended to in programs for the intellectually gifted (Houtz et al., 1978, p. 517).

Sherman (1977) used three tasks which sampled different steps in the CPS process: preparation, brainstorming of problems in the fuzzy situation; idea production, or solution finding; and evaluation, a judgment task in which good ideas or poor ideas were generated. All three tasks were phrased to elicit responses where they could be scored for ideational fluency. Intercorrelations of the fluency scores for each CPS task, self-esteem, locus of control, IQ, and four achievement variables were computed. The results provided partial support for the hypotheses for the study. When data were analyzed for the combined grade (fourth through sixth) groups, the two affective variables, self-esteem and locus of control each showed a relationship to achievement, and locus of control was also related to the preparation phase measure of fluency. However, when the data were analyzed separately for each grade level, relationships were not found between locus of control and any of the CPS fluency tasks. Achievement and intelligence supported hypothesized relationships, as did self-esteem and achievement, and self-esteem and locus of control. Little evidence was found to support a relationship between intelligence and CPS ideational fluency nor were relationships established with self-esteem or locus of control. Sherman's data lend support to findings of a relationship between locus of control and achievement (Crandall, 1978); and between self-esteem and achievement (Sears, 1972). It must be noted that in these studies the children were tested on tasks of CPS processess but had not received any training in CPS method to facilitate creativity. It would be interesting to determine if students trained in the CPS methods might show a clearer relationship of the affective measures and effectiveness of their skills in CPS, and, if so, whether such skill would vary with the different phases of the CPS process.

Affective Characteristics, Creativity, and the Gifted

In this section, the need for research using affective measurement specific to creativity is introduced. Studies of gifted childrens' self-concept and of the relationship of self-concept, locus of control, and creativity are surveyed, and those which relate to creativity, self-concept, locus of control, and the gifted are discussed. Studies which bear on the selection and/or development of the three affective instruments used in this study and their validity are included in the review.

General and Specific Self-Concept

Self-concept refers to the individual's cognitive picture of the self (Whitmore, 1980). It covers the broad aspects of how the individual views who or what he is in regard to traits of personality, ability, and appearance. From the person's perceptions of relative strengths or weaknesses in that picture, expectations of relative success in most activities are likely to be generated. The self-concept, as the core around which all personality is organized, then, can have a profound influence on the person's behavior by affecting how well a person expects to succeed in various aspects of life (Whitmore, 1980). The question, however, with use of the construct of general self-concept, is that if the construct covers all aspects of the person's cognitive picture of the self, would it then be reasonable to expect relative success or failure in a specific area of behavior to change an inclusive picture of the self? The following argument develops reasoning for the use of specific self-concept measures to examine relationships with expectations to relevant areas of behavior.

The perceptions contributing to self-concept are formed through experiences of personal environment and interpretation and reinforcements of the meaning of those experiences (Shavelson & Bolus, 1982). A structure of self-concept as hierarchical and multi-faceted was proposed by Shavelson, Hubner and Stanton (1976). Shavelson et al. (1976) viewed academic self-concept as separate from a construct of non-academic self-concept consisting of social, emotional, or physical subareas of the self-concept. Winne, Marx and Taylor (1977) found some support for the idea that individual facets of self-concept were related in varying degrees to other constructs. However, Winne et al. viewed self-concept as basically a unitary, undifferentiable construct.

Shavelson and Bolus (1982) reexamined the data in the multitrait-multimethod study of self-concept by Marx and Winne (1978). The Marx and Winne (1978) data from the three self-concept inventories of Sears, Gordon, and of Piers-Harris, were restructured by Shavelson and Bolus into a new multitrait-multimethod matrix which examined three subareas of the self-concept: physical, social, and academic. In the restructured analysis, evidence supported the multifaceted view of self-concept. A strong correlation existed within methods (instruments), however, "different measures of the same trait are more highly correlated with another than with different measures of different traits" (Shavelson & Bolus, 1982, p. 5). A direct implication of Shavelson and Bolus' work is that a multifaceted, hierarchical interpretation of the construct of self-concept is highly useful for research as well as theoretical purposes.

A test of the assumptions of the hierarchical construct required measurement of two or more levels of the hierarchy in at least one area or facet of the general construct (Shavelson & Bolus, 1982). The area of academic self-concept was selected in a pretest-posttest investigation of two measures of global self-concept, two measures of academic self-concept, and two measures of self-concept specific to subject matter areas. Analysis of the covariance structure of the data failed to support the hypothesis of differing degrees of stability; general self-concept was not more stable than specific self-concept.

However, because there was no treatment condition in the Shavelson and Bolus (1982) study, it is believed that a design which would contain a treatment condition in one of the specific subareas could offer a better test for the hypothesis of increasing stability toward the apex of the hierarchy, i.e., the concept of self that is the most general.

Other than the question of differential stability, the data supported the model of multifaceted, hierarchical structure for the self-concept. General self-concept was found to be distinct from, but related to, academic self-concept. The specific areas of subject matter self-concept were found to be intercorrelated but distinguishable from each other. Achievement in the respective academic areas, as measured by grade, bore a close correlation to the specific relevant academic self-concept and some relation to general academic self-concept but little relation to general self-concept. Moreover, the specific academic self-concept showed a correlation with the measure of academic self-concept and with the measures of general self-concept. The authors applied three different models of self-concept to the covariance structure and found that the most differentiated, "full, multifaceted model accounted for 80% of the covariation" (Shavelson & Bolus, 1982, p. 11).

The investigation also examined causal paths between differing measures of self-concept and each achievement measure. An interesting finding was a causal path from grades in an English pretest to a slightly negative coefficient to the general self-concept posttest rather than to a positive correlation with the subarea academic or specific (English) self-concept. The data suggest an interpretation that the "negative relation between general self-concept and grades is probably due, then, to the nonacademic facets of general self-concept such as social self-concept" (Shavelson & Bolus, 1982, p. 14). Cognizance of the possibilities of negative or no correlation between a specific achievement area and a general measure of self-concept is possibly important in interpreting research in the area of general self-concept and creative behaviors of the gifted where predicted relationships have failed to be exhibited in the data (Fults, 1981; Kolloff & Feldhusen, 1984).

Other authors have explored areas of self-concept specific and relevant to the content being studied. A study of use of subscales rather than total score of the Piers-Harris Children's Self-Concept Scale obtained dramatic differences with the use of subareas of self-concept to meaningful hypothesized relationships for sex, race, age, or social class groupings (Osborne & LeGette, 1982). In utilizing subareas of the Piers-Harris Children's Self-Concept, Kanoy, Johnson and Kanoy (1980) were able to locate meaningful differences between achieving and underachieving bright children on the subarea related to intellectual and school status. Williams (1976) suggested that research in the area of creativity and self-concept might profit from differentiating between an academic self-concept and personal self-concept.

Self-Concept Specific to Creativity

An assessment of literature by Wright, Fox and Noppe (1975) on the relationship of creativity and self-concept revealed contradictory findings. The results of several studies cited by Wright et al. suggested that groups of persons high in self-esteem as measured by the Coopersmith Self-Esteem Inventory (Coopersmith, 1967) were high in product measures of creative thinking, and persons low in self-esteem tended to be consistently less creative. The persons with greater self confidence showed greater creativity by their willingness to seek novel solutions to problems.

These results are consistent with Maslow's (1962) description of self-actualizing or fully-functioning, self-accepting persons as creative. A continuum analogous to Maslow's position was provided by Sund and Carin (1978):

participation in creative enterprise builds capabilities, increases positive self-concept, makes people more open and free, and thus contributes to good mental health.

On the other hand, Arieti (1976) described a commonality of creativity with neurosis, as seen by Freud, who hypothesized that both originate in conflicts. Wright et al. (1975) located other studies which did not find a significant relation of self-regard to creative potential as measured by the Revised Art Scale of the Welsh Figure Preference Tests, or to creative products as measured by some of the Guilford tests.

The foregoing apparent conflicts make sense in light of findings that highly creative productive persons exhibit a unique blend of opposite qualities (Torrance, 1979). They are more anxious, disturbed, and discontented, yet psychologically healthier than the average person. The blend of energy and ability to synthesize in a unique manner is seen as the creative "delicate balance" between deferred judgment and judgment (Parnes et al., 1977); the "magic synthesis" (Arieti, 1976); and the "unity of opposites" (Torrance, 1979, p. 5). Wright et al. (1975) recognized the need for a self-concept scale specific to creativity. A description of their scale and findings is discussed below under measurement of specific self-concept.

Measurement of Self-Concept

Measurement of Self-Concept in Studies of Creativity Training with Gifted Students

As can be seen from the preceding analysis, it may be difficult to establish a hypothesized relationship of positive change in general self-concept due to programming in a specific area. General self-concept, usually comprised of subareas of personal, academic, and social self-concept (Shavelson & Bolus, 1982) or even, hopefully social, academic, and divergent self-concept (P. S. Sears, personal communication, July, 1985, see Appendix A) is likely to be stable and resistant to change due to treatment in a specific area such as creativity training (Wright et al., 1975; Shavelson & Bolus, 1982). Several recent studies utilized the Piers-Harris Children's Self-Concept Scale (PHCSC) as the general self-concept measure of gifted children in a special program, which included some emphasis on creativity training. Significant gains were not shown in PHCSC general self-concept from pretest to posttest as associated with creativity training (Fults, 1981; Jones, 1983) or in a posttest only design with an experimental and control group of gifted students (Kolloff & Feldhusen, 1984).

Tadlock (1981) examined component scores of the PHCSC rather than total scores and TTCT. She failed to find significant differences in any PHCSC score or in TTCT scores between the second grade gifted treatment or gifted control group. Third grade students, however, in the treatment condition of basic movement education, showed significant differences in the TTCT components of fluency, originality, and elaboration, and one of the six PHCSC components, that of behavior, as a contributor to positive self-concept (Tadlock, 1981).

One might at this point question whether perhaps the PHCSC could be less sensitive when measuring changes in a sample of children of a restricted IQ range, i.e., the gifted. However, in a study of achievement study skills, in an academically able but underachieving group (grades six through nine), PHCSC scores were raised (Crittenden, Kaplan & Heim, 1984). A different interpretation, then, could be that the PHCSC is not an instrument sensitive to the types of changes in self-concept that might be expected when treatment is in a creative, rather than an academic, area. Studies which used a single PHCSC score rather than a repeated-measures design to compare gifted children in different program designs found differences between groups (McQuilkin, 1980/1981) or between gifted children who describe themselves as feeling "different" in a positive fashion and gifted children who do not think of themselves as different (Janos, Fung & Robinson, 1985). In an analysis of differences in program options where the repeated measures were observed three times over a period of 18 months, significant findings were again noted (Coleman & Fults, 1982). A more meaningful interpretation, then, might be that a measure of total self-concept which has content items suitable for assessing creative self-concept and/or a separate measure of creative self-concept might be a more appropriate measure than the PHCSC for studying affective aspects of change in programs of creativity training.

General Self-Concept: Sears Self-Concept Inventory

A general self-concept scale for children, which recognized assessment of feelings about oneself as a creative person as contributing to the total picture of the self, was developed by Sears (1963, 1975). The nine areas of self-concept were determined by a priori judgment of items which ought to be coherent, one with another (Sears, 1975).

Sears reported two subsequent analyses which provided empirical confirmation of a divergent self-concept subscale (P. S. Sears personal communication, July, 1985, see Appendix A) by means of two factor analyses of the Sears Inventory (Sears). She reported three strong factors of 35 of the 48 items with loadings greater than .30 and labeled them as Academic, Social, and Divergent self-concept. Three questions loaded on both Academic and Divergent factors. Ewert (cited in P. S. Sears, personal communication, July, 1985) reported eight factors for 40 of 52 items at greater than .40 (four new items were added to specify "social relations with own sex" as ". . . with boys" and ". . . with girls"). Ewert labeled the factors as: Creativity, empathy; Outer appearance (physical), ability of writing; Physical abilities; (not named, but could be named as "motivation"); Social relations with boys . . . and with girls; Leadership, dominance;

Convergent mental; Autonomy, self-confidence; and School adjustment. The factor analyses provided evidence toward construct validity that the full self-concept scale contains a creative portion.

Criterion validity studies of the Sears were conducted as part of a five-year study of effective reinforcement behaviors for disadvantaged children (Sears, 1972). Pre-and posttest data during the first, non-interventional year were analyzed to aid in selection of the various instruments and to examine relationships between behaviors typical of the teachers' and children's achievement and attitudes. A step-wise regression analysis of variables to predict posttest scores of the full self-concept score located the following predictors: pretest scores on self-concept, $r = .52$; child social distance self-rating, $r = .35$; teacher rating of the child in the social distance scale, $r = .17$; two variables of teacher behavior, "criticize with explanation" $r = -.23$ and "private approval to individual" $r = -.15$; and Hess locus of control, I-, $r = .04$. The prediction of self-concept was found to be less accurate by the regression equations than the prediction of achievement. The final prediction for posttest achievement accounted for 80 percent of the variance and the prediction for posttest self-concept accounted for about 54 percent of the variance (Sears, 1972).

A study investigated differences in academic and social self-concept of the academically gifted utilizing the Sears (Ross & Parker, 1980). The authors found a significantly lower social than academic self-concept among the fifth through eighth-grade gifted students in the sample. Ross and Parker's results, however, should be interpreted most cautiously because no corrections were made for differences in length of the respective scales (Ross & Parker, 1980; Sears, 1975). That is, when the Sears test was divided into two subscales, social and academic, there were fewer questions on the social scale.

Specific Self-Concept: Creative Self-Concept

A brief ten-item Likert-type scale for creative self-concept was designed by Wright et al. (1975). Wright's scale and a general measure of self-concept, the Tennessee Self-Concept Scale, along with two measures of creativity, a verbal and a figural subtest of the TTCT, were administered to college students. Intercorrelations among the eight scores of measured creativity and the two self-concept measures yielded significant correlations of the creative self-concept with three verbal measures of the eight creativity measures. General self-concept failed to approach a significant relationship to any of the eight creativity scores except verbal elaboration. The specific creative self-concept bore a significant but relatively low positive correlation to the general self-concept. Two multiple correlations were computed to test the relationship of each of the self-concept variables to measured creativity. A significant multiple correlation of measured creativity and creative self-concept as a criterion variable ($R=.44$, $p < .05$) was found. However, a correlation was not found between the TTCT scores and general self-concept as hypothesized, thus showing evidence for construct validation of Wright's scale.

The lack of relationship between creativity scores and general self-concept in light of the relationship between the specific creative self-concept and measured creativity was surprising to Wright et al. (1975). They concluded that "while college students who perceive themselves to be more creative are more creative . . . a global approach to self-concept appears not to be related to objective measures of creativity" (p. 13).

The findings of Wright et al. (1975) fit with the consistent finding of the preceding review; there is a failure to establish change of total self-concept scores when measured creativity increases. The investigation by Wright establishes criterion validity of a creative self-concept scale as related to measures of creativity and of marginal relationship to general self-concept. The authors recommended incorporation of a subscale of creative self-concept in development of future scales of general self-concept.

Locus of Control, Creativity, and the Gifted

In this section, the construct of locus of control as a theoretical basis for interpreting behavior is briefly introduced. Studies which bear on locus of control in gifted children, or on the relationship of creativity, locus of control, and the gifted will be discussed. Prior research on validity of the proposed locus of control instrument, the Intellectual Achievement Responsibility Scale (IAR), will be examined, especially that which pertains to the gifted and/or creativity and demonstrates the sensitivity of the IAR as a measure of treatment effects.

Generalized and Specific Expectancies of Internal Versus External Control of Reinforcement

The construct of internal-external control of reinforcement was theoretically proposed and experimentally reviewed by Rotter (1966/1982) as a unit within his social learning theory. In the 1982 introduction to the volume of his collected papers, Rotter described social learning theory both as a process and content theory and as an interactionist approach to personality, where behaviors are viewed in relation to each other and the environment. A major contribution of social learning theory is its attempt to "integrate two diverse but significant trends in American psychology--the stimulus-response, or reinforcement, theories on the one hand and the cognitive, or field, theories on the other" (Rotter, 1975/1982, p. 267). Some of the basic assumptions of the theory are that personality has unity, in that a person's experiences with the meaningful environment influence each other; that behavior may be described as goal-directed with the directional aspects of goals inferred from reinforcers; and that behavior is influenced by expectancies in regard to goals or reinforcers. The expectations are derived from a person's prior experience and may be measured utilizing operationally definable constructs and empirically testable hypotheses (Rotter, 1982).

An individual's perception of the relation of reinforcers in his personal environment and in his own personal behavior can be seen as expectancy of the power to exert influence of one's environment to the individual's life. Persons who perceive that outcomes occur mainly because of their own effort or ability are considered to have an internal locus of control, and those who perceive events in their lives as due to fate, luck, or external work of powerful other persons or other forces are considered to have an external locus of control (Bradley & Gaa, 1977; Rotter, 1982).

Some problems in the research literature on internal-external control were summarized by Rotter (1975/1982) as failure to account systematically for the value of the reinforcer by control or measurement of it as a separate variable. The specificity versus the generality of the expectancy is a problem area in research pertaining to achievement. Some successful prediction of achievement, with ability as a controlled variable, may be made as a function of attitudes toward locus of control in the early grades, but the prediction is less pronounced for older students. It appears that although there is

> . . . a persistent effort to obtain highly accurate and reliable predictions of achievement behavior . . . it becomes less reasonable the more structured, the more familiar, and the more unambiguous a particular situation is (Rotter, 1975/1982, p. 271).

A greater degree of prediction occurs between learning, and achievement-related, variables when the material is relevant to the person's goal strivings. Rotter stated that another major problem in internal-external research is the tendency for some psychologists to assume that it is good to be internal and bad to be external. Rotter considers this as oversimplifying the complexity of relationships in adjustment of the individual.

<u>Expectations of control specific to achievement</u>. An investigation into the generality-specificity issue utilized two specific locus of control instruments: the Intellectual Achievement Responsiblity Scale (IAR) and the Locus of Control Inventory for Three Achievement Domains (LOCITAD) (Bradley & Gaa, 1977). Individual goal-setting conferences to assess classroom work, evaluation, and progress toward previously set goals were used for five weeks. Posttest scores for locus of control of the tenth grade English students were compared to those of students who had individual scheduled weekly conferences that were not of a goal-setting nature and to those of students in the same classes who received no special conferences. A multivariate analysis of covariance with subject matter achievement as a covariate showed significantly different effects between the goal-setting treatment groups and those groups that did not meet to set academic goals for the IAR scales (internal responsibility for positive events, I+, or for negative events, I-), and LOCITAD (intellectual) scores but not on the LOCITAD scores specific to the social or physical domains. Other comparisons between the groups were consistent with the conclusion that the goal-setting conference produced more internal expectancy of control. The impact was limited to the expectancies of control in academic situations, without generalizing to other types of situations. Bradley and Gaa (1977) suggested further that studies using these specific measures would yield information not obtainable with more generalized expectancy of control measures. They suggested that a person's control orientation to a rather specific class of situations might be changed and need not interfere with that person's ability to maintain orientations of control appropriate to other given situations. Bradley and Gaa's findings support Rotter's (1975/1982) position that researchers should be careful in making generalizations

that an internal locus of control is without qualification the preferred orientation for all situations.

An attempt to generalize effects of increased internality with respect to achievement orientation found that the behavioral task used was unable to show discrimination between internals and externals (Companik, 1978). Experiences with task success trials assisted in providing IAR externally oriented children with a more internal orientation. However, responses to the experimental task did not differentiate between internal and external individuals. It is possible that the interpretation of specificity of locus of control, as proposed by Bradley and Gaa (1977), would appropriately account for the increased internality with respect to achievement orientation to generalize toward internality in the social domain.

Differences between sexes have been found for locus of control expectancies to achievement. Successful achievement related prediction seems to occur primarily for locus of control in boys (Crandall, 1978; Lefcourt, 1966).

Locus of Control Expectancies and the Gifted

The use of specific measures showed a difference in intellectual and school status scales of the Piers-Harris Children's Self-Concept scale and of the IAR total, I+, and I- scales between 20 achieving and 9 underachieving bright fourth-grade students (Kanoy, Johnson & Kanoy, 1980). They found achievers to be higher than underachievers in self-concept for intellectual and school status, in IAR total score, and in I-. These findings support the reasoning by Crandall (1978) that "internal-external perceptions might not be equivalent because it's more difficult to assume responsibility for failures than successes" (Crandall, 1978, p. 1). Unlike researchers with other populations of children, Kanoy et al. (1980) did not find sex differences in locus of control or academic self-concept related to achievement in their sample of bright children between those who were achievers and those who were underachievers. A difference in self-concept on the subscale of popularity showed an interaction between sex and achievement which was not readily interpretable. Because Kanoy et al. used several separate two-factor and single-factor ANOVAs rather than a multivariate approach to examine these data from a small sample, a not readily explainable sex times achievement interaction could be an example of a chance effect due to a possible problem of an inflated α level in this study.

Several tasks representing the hypothesis-generating (divergent) phase and the hypothesis-testing (evaluative) phase of the CPS process were presented to elementary gifted children along with three affective measures: tolerance of ambiguity, self-esteem, and locus of control (the Bialer-Cromwell Locus of Control Scale (Tetenbaum & Houtz, 1978)). This study explored the relationships between the affective traits and the different tasks representing phases of the creativity and problem solving processes. The results of the analysis pertaining to the cognitive creativity and problem solving tasks were described above.

Studies using the IAR usually find girls to score higher than boys, particularly at older age levels (Crandall, 1978) and in a sample of intellectually superior early grade students (Crandall, Katkovsky & Preston, 1962). In these above cited studies, however, no differences between bright or gifted boys and girls were found with respect to locus of control (Kanoy et al., 1980; Tetenbaum & Houtz, 1978) nor the other affective measures. Grade differences on the affective measures were found (Tetenbaum & Houtz, 1978) only on tolerance of ambiguity, showing fourth-grade students to be less tolerant than those in either fifth or sixth grade.

Tetenbaum and Houtz (1978) demonstrated a relationship between their affective and cognitive measures by analyzing the data with a canonical correlation. The composite showed that the affective and cognitive variables accounted for about 46% of the variance in the first canonical correlation. A correlation of each of the original variables with the first canonical variate showed that the following variables contributed to the canonical solution with a load $\geq .30$: tolerance of ambiguity and locus of control in the affective set, and ideational fluency, expressional fluency, patterns and numerical reasoning in the creative set. The predictions of a relationship between locus of control and the cognitive problem solving variables and another separate relationship between tolerance of ambiguity and the creativity measures were not confirmed as separate sets of correlated dimensions. Nevertheless, the Tetenbaum and Houtz study showed that affective measures share a large portion of the variance with measures of tasks related to CPS in gifted students and, therefore, are important considerations in understanding creative behavior.

Other studies from Fordham University investigated locus of control as one of the affective characteristics considered in relation to tasks of creative thinking (Beck, 1979; Houtz et al., 1978; Sherman, 1977). Houtz et al. found that when variability due to individual student achievement on standardized tests was removed, locus of control and self-esteem, but not tolerance of ambiguity, were indeed significantly related to performance or the cognitive measures of creativity and problem solving.

> One might, therefore, conclude that the characteristic of high tolerance for ambiguity is most important in achievement situations while an internal locus of control and high level of self-esteem are more important in creative thinking situations above and beyond achievement demands (Houtz et al., 1978, p. 517).

In her review of the literature, Sherman (1977) found that "very little research has been done which relates the locus of control construct to creative ability and problem-solving ability" (p. 40). Sherman's discussion proposed that the characteristics of persons who are internally controlled parallel those found in studies of creative individuals. Internally controlled individuals may be described as differing from externals in that the former are:

> . . . more autonomous and independent in behavior, active in dealing with the environment, resistant to external influence where not appropriate, higher in intellectual efficiency, effective and industrious, less anxious and demonstrating more initiative in his efforts to attain goals and to control his environment (Sherman, 1977, p. 43).

Several parallels to the present author's analysis of the literature of traits of creative persons are found to Sherman's analyses of traits of internal individuals. Creative individuals may be considered to possess self-respect and good sense, honesty of thought and behavior (Barron, 1963), and to be courageous and pre-occupied (MacKinnon, 1975). Independence of judgment and thought was found in studies by Barron (1963) and MacKinnon (1975) and was a recurrent characteristic noted in Arieti's (1976) discussion of specific personality characteristics of the creative individual. Heist (1968) noted that college students considered by their instructors to be creative differed from those identified as scholars; creative students were more judgmental and scored high on scales of Autonomy and Religious Liberalism. Of nineteen characteristics of the creative individual cited by Stein (1974-1975) in his summary review of the literature, the following especially parallel those of the internally controlled individual: self-assertiveness, independence, willingness to take initiative, persistence in motivation, rejection of repression, less concern about conventional standards, and determination. Some of the traits of persons considered to be highly creative, then, are parallel to traits of the internally controlled individual.

In studying locus of control and self-esteem as personality characteristics in relationship to divergent thinking performance on tasks related to CPS, Sherman's (1977) findings were consistent with those of Tetenbaum and Houtz (1978) that internal locus of control is related to tasks requiring fluency of thought. Sherman found locus of control to be related to ideational fluency on tasks related to the preparation and production of ideas phases of CPS but not to the judgment phase (producing good ideas and bad ideas). Sherman's data with gifted elementary children confirmed the findings of other studies (Crandall, 1978); that internality tends to increase with age.

In a study to investigate possible effects of labeling on students prior to receiving service in a gifted program, Gabrielle (1985) used a pre-post design to test 75 high achieving ninth graders. An increase in scores for locus of control for negative events (I-) and for total locus of control (I_t) was also found to occur in ninth graders as a concomitant effect of identification or labeling for a gifted program prior to receiving service in the program (Gabrielle, 1985). All of the students who were identified as gifted felt positive about being selected for the program.

Studies of Locus of Control and Creativity in Other Populations

Community college students were the subjects in a study by Beck (1979) in which groups of subjects differing in locus of control orientation were compared in their performance on an ideational fluency task administered under two types of instruction. The findings were generally supportive of the premise that type of instructions to a task can moderate the effect of locus of control on CPS performance. Three sets of instructions were used: all subjects first performed the CPS tasks under instructions designed to increase the role of chance and were later retested with instructions which emphasized the role of skill, but in half the subjects, the threat of external evaluation was minimized, and in the other half, it was maximized. The chance instructions minimized differences between locus of control and CPS performance. However, under the retesting condition of manipulating threat of external evaluation, a complex interaction was found. The results suggest that externally oriented individuals surpassed those with a chance or with an internal orientation under non-evaluative conditions, but performed less well than internal subjects under evaluative conditions (Beck, 1979).

A study to determine the relationship between affective characteristics of locus of control, test anxiety, and field independence, and performance on convergent and divergent thinking tasks was conducted by Elkind (1969). The three affective measures were found to be essentially independent descriptions of the fifth grade students in the sample. The combination of characteristics considered as descriptive of intrinsically oriented children--low test anxiety, internal locus of control (using the IAR), and field independence--was significantly associated with higher performance on the convergent tasks for the intrinsically than externally oriented subjects. Test anxiety was unrelated to the divergent tasks. Internal locus of control and field independence were associated with higher performance on the TTCT verbal tasks but were unrelated to the performance on the TTCT figural task. It is possible that the use of fairly distinct racial and socioeconomic groups in the sample could have a confounding effect on interpreting these findings in relation to divergent thinking and intrinsic versus extrinsic orientation. The black students from the lower socioeconomic status school were more extrinsicially oriented but performed better than the other group in the non-verbal divergent task. It is not, however, a surprising finding that children from a lower socioeconomic status will show a figural-verbal discrepancy in favor of figural TTCT (Torrance, 1974b).

Intellectual Achievement Responsiblity
and Specific Locus of Control

A shortened form of the IAR was developed with use of a new normative population of 1,793 children (Crandall, 1978). It was necessary to discard the norms obtained from the original normative sample (Crandall et al., 1965) because an historical shift was observed to occur toward externality of responses over the years.

A factor analysis was performed on the current data for the I+ scale and separately for the I- scale. Seven pairs of items with the highest factor loadings, all > .51, were selected for the short form. The content of six of the seven pairs was entirely school-related. Construct validity was higher with use of the short form than the full form because "five of the seven pairs are effort items, which is really what locus of control is all about" (Crandall, 1978, p. 7). The items of either the full form or the short form are relatively free from social desirability bias (Crandall, 1978).

Persistence of Effort in Relation to Locus of Control

In a review of research developments with the IAR, Crandall noted that the IAR "has proved helpful in predicting task persistence and effort" (1978, p. 2). One of these studies will be briefly discussed to demonstrate the relationship of internal locus of control as measured by the IAR to task persistence and as evidence of sensitivity of the IAR to treatment effects.

From children in upper elementary grades who were identified as well below grade level in reading, 28 were selected as the most helpless on the basis of low IAR scores and performance on an effort versus ability failure scale (Fowler & Peterson, 1981). Teacher ratings confirmed the validity of the selection process for children exhibiting helplessness with respect to lack of self-confidence and academic task persistence. The children were assigned to one of four different treatment groups for a three-day training period. The treatment conditions consisted of two schedules of reinforcement in which the child would experience success with 16 difficult sentences. Two feedback conditions were added to the reinforcement schedule which required greater persistence, in order to vary indirect or direct attribution of success or failure to personal effort. The examiner under this condition stated, "No, you didn't get that. That means you have to try harder" (Fowler & Peterson, 1981, p. 252). A direction attribution of responsibility reinforcement technique required the children to verbalize a similar reinforcement, aloud, in a whisper, and silently.

A significant treatment effect was found on posttest scores for improving persistence at attempting a greater number of sentences as an interaction between reinforcement treatment groups and time, with children in the indirect and direct attribution of responsibility conditions showing greatest improvement. Differences between IAR pre- and posttest scores were found after three days of training in the groups of children who received the training using direct attribution of effort (verbalization condition). The results showed on the IAR effort scales for I+ and I-, but did not change the IAR_t score.

A clear implication of these studies with learned helpless children is that remedial programs which aim toward removing failure experiences may do a disservice to the children by rendering "children incapable of dealing with subsequent failure that they will inevitably encounter" (Fowler & Peterson, 1981, p. 259). The results obtained by Fowler and Peterson (1981) are consistent with the above analysis of the literature recommending that a specific measure of locus of control is likely to be sensitive to experimental differences. It also provides support for criterion-related validity of the IAR effort scale.

Attribution of responsibility for one's actions has been shown to relate to school achievement. Moreover, scores on the IAR effort scale increased after training to exert greater effort on a task. Evidence was presented to indicate that a relationship exists between internality of locus of control and verbal fluency as well as performance of creative thinking tasks.

Social Aspects in Group or Team Creativity Tasks

In this portion of the review of the literature, a brief review is presented of the social aspects of performance in relation to creativity tasks and to creativity. This section is presented at the awareness level to introduce a variety of factors which may have some degree of relevance to the behaviors investigated in this study.

Creativity Phenomena in Groups

Amabile (1983) noted that prior studies of creativity have been largely concerned with the study of personality characteristics of creative individuals, descriptions of persons who do well on creativity tests and those who do not, or studies of the creative process by cognitive psychologists. She proposed the need for "a social psychology of creativity {which} aims to identify particular social and environmental conditions that can positively or negatively influence the creativity of most individuals" (Amabile, 1983, p. 5). The perspective of the social psychologist permits examination of the role of source of data, such as first person reporting and other perspectives discussed below.

Amabile (1983) advocated the consensual assessment of creativity by two or more judges utilizing implicit criteria of the creative product, as well as assessing some dimensions such as technical aspects of the creative product, and the rating of such products relative to one another rather than to an absolute criterion (Amabile, 1983). It is interesting to note that the scoring methods used for OM competition fit some of the methods involving consensual judgement (Micklus & Micklus, 1987; Purifico & Micklus, 1987). Judges' individual scores of a given judging cateogry result in an averaged score, and where there are important differences of opinion among members of the judging team, it is suggested that consensual agreement should be reached. However, due to competition constraints, the different judges are unable to view the products in different orders of presentation, a procedure suggested by Amabile.

Some of the research examined by Amabile (1983) surveyed the impact on creative performance of intrinsic versus extrinsic success of individual motivation and the deleterious effect of reinforcement on creative production. She also reviewed effects of performance in the presence of an evaluative audience, importance of influential models on creative individuals, and facilitating and inhibiting effects of the mere presence of others.

Is Creativity Facilitated by Membership in a Group?

> Social facilitation research has shown that the mere presence of others--either as coactors or as an audience--can impair performance on poorly learned or complex tasks, but enhance performance on well-learned or simple tasks. Most of the evidence also suggests, however, that in humans the expectation of evaluation can augment these social facilitation effects (Amabile, 1983, p. 141).

However, in only two of these studies was the dependent variable a task requiring creative performance. In a study of a word association task, performed either alone or in the presence of others, subjects gave quicker responses, but more common responses, in the observed condition (Matlin & Zajonc, 1968). Normal children performed less well in individual creativity tests when tested in the presence of others, but no difference was found for gifted children between individual and group administered scores (Milgram & Milgram, 1976).

The social facilitation experiments for the most part did not involve complex behaviors nor ones which produced divergent responses. The Cottrell (1968) research tended to support the interpretation that the presence of others is a source of drive. Such drive may increase performance of well-learned behaviors. The presence of others in a brainstorming or CPS group is considered to be motivating (Osborn, 1957).

Much of the social facilitation research, it should be noted, differs from the process of brainstorming in a group such as in CPS work where individuals work with each other rather than merely in the presence of another person acting as an audience or a passive observer (Amabile, 1983; Matlin & Zajonc, 1968; Osborn, 1957).

The effect of the presence of others in a passive or active role is difficult to ascertain from psychological studies of creative individuals. A sociological study was considered to provide "only suggestive evidence for the facilitative effects of other people on individual creativity" (Amabile, 1983, p. 142). In a study of 286 Nobel prize winners from 1901 to 1972, it was found that nearly two-thirds of the Nobel Laureates had researched their outstanding work in conjunction with others (Zuckerman, 1977). The trend toward collaborative work is increasing in all of the sciences, typically in teams of two or three, with some recent evidence of larger groups working together. A direct implication is that work done in a collaborative team provides a situation which can facilitate original and successful creative work.

Teamwork which is experienced as collaborative is conducive to increasing creativity (Osborn, 1957; Zuckerman, 1977). "Most of us can work better creatively when teamed up with the right partner because collaborativity tends to induce effort and also to spur our automatic power of association" (Osborn, 1957, p. 72). Osborn attests to the productivity of group brainstorming sessions in terms of social facilitation effects and of stimulative effect of rivalry. "On the other hand, if it is true that the presence of others leads to increased arousal and, perhaps, to evaluation expectations, then creative performance should suffer under such conditions" (Amabile, 1983, p. 142).

If a student, then, perceives the experience as one that is collaborative and feels that other members of the team as well as he participated fully, exerted effort, communicated well to produce ideas, and produced good work, then increases in creativity and in internality of the effort scale of the IAR would be expected. If, however, negative anticipations of evaluative judgments are perceived by the student of his work (Cottrell, 1968), particularly on CPS tasks which are indeed complex, difficult, and unfamiliar (Matlin & Zajonc, 1968), then the effect might be felt as evaluative working in the presence of others rather than working with collaborative others and would be experienced as undermining to the child's creativity on CPS tasks.

Evidence to the contrary, however, was suggested by Amabile (1983) from studies where the subjects perform more poorly in idea-production tests when they work together than when they work alone. "Overwhelming evidence" (Amabile, 1983, p. 143) from several studies (Taylor et al., 1958 cited in Stein, 1974-1975; Renzulli, Owen & Callahan, 1974 cited in Amabile, 1983) shows that people are more productive when working alone in nominal groups rather than when they work together in a real group. The ideas produced were judged by Taylor in terms of fluency and quality (Taylor, 1958) as measured by "uniqueness," "effectiveness," "feasibility," and "generality."

A closer reading of the studies reported in Stein (1974-1975) does not find the evidence to be overwhelmingly in favor of individual work as suggested by Amabile (1983). In a further analysis in which Taylor (1958, cited in Stein, 1974-1975) controlled fluency it was found that the only difference was that nominal groups produce more ideas than real groups. The reanalysis did not support the earlier findings of higher quality of responses. A later study of Bouchard (1969, cited in Stein, 1974-1975) suggests that the differences between Taylor's groups could be an artifact of the manner of collecting the data. It takes a person longer to write down ideas than to speak them, and four persons speaking alone can say more ideas in a 12-minute time limit than can four persons working individually who are later combined into a "nominal" group.

Amabile (1983) considered scores on the most popular tests of creativity, such as the TTCT, to primarily reflect the processes of an algorithmic task, where "the path to solution is clear and straightforward and responses to it simply cannot be considered creative" (p. 101). Tests which yield scores of fluency and flexibility, according to Amabile, are algorithmic when the algorithm for solution is known. Essential to her definition is that the "creative response is a novel and appropriate solution to a heuristic task . . . the task must be open-ended to some degree" (Amabile, 1983, p. 101). Amabile (1983) considered studies of competition and creativity such as those conducted by Torrance in his work in the 1960s. Torrance's studies showed higher scores on fluency and flexibility under conditions of competition and/or reward. Based on her conceptual position and review of the literature, however, Amabile concluded that extrinsic motivation, such as the reward of competition, might facilitate performance on algorithmic tasks but would be likely to impair performances on heuristic tasks.

The classification of tasks as to algorithmic or heuristic is in part dependent on whether the algorithm is known to the learner (Amabile, 1983). Perhaps responses to the TTCT could be considered to a limited extent to be algorithmic in that the TTCT is subject to training effects (Rose & Lin, 1984). Based on the present author's review of the literature and study of the tasks, it is most difficult to classify the tasks on the TTCT as basically algorithmic and not truly open-ended.

In a review of the literature on competition and creativity, Roweton (1982) concluded that "even after listing all these concerns, the data stands resolutely. So, it still must be admitted that competition facilitates creativity, at leas {sic} as measured here" (Roweton, 1982, p. 94). Roweton, in closing, commented that organizations could profit from being sensitive to the best conditions for different individuals and different tasks. Perhaps the conflicting qualities of some of the research on groups and competition might be accounted for by accounting for some of the variability of subjects in CPS groups.

In a study of brainstorming sessions, college students working in cooperative dyadic pairs scored higher on originality and flexibility and somewhat higher on fluency than those who worked individually (Bolen & Torrance, 1978). The MANOVA of the two TTCT activities also showed the males to be more flexible than females. Unlike the results from related research, subjects with an external locus of control were found to be more flexible than those with an internal or mixed orientation for responsibility.

In summary, evidence was reviewed by Amabile (1983) suggesting that creative thinking is inhibited by performance in a group. The affective effects of working as a group on a creative task were not discussed. An alternative hypothesis can be offered based in part on Amabile's review and that of Stein (1974-1975). Although individuals working in a nominal group rather than a real group may appear to do better, evidence is presented that individuals can indeed perform better on creative tasks when they experience the work in their group as cooperative (Bolen & Torrance, 1978; Osborn, 1957; Roweton, 1982; Zuckerman, 1977). Examination of the variable of effort in creative work by groups could possibly account for some of the conflicting findings of beneficial or deleterious effects on creativity by individuals who work in a group.

Effort and Creativity

In this section, effort is described in relation to task commitment and to creative productivity. Effort, persistence, and gifted behavior are discussed in relation to academic achievement. Some operational definitions of effort are surveyed, and self-efficacy and effort are briefly discussed in relation to CPS.

Application of effort was noted by Osborn (1957) as a necessary condition for true creativity to occur. He stated that improved concentration intensifies awareness of the relationship among ideas and increases the likelihood of creating more fruitful associations.

Persistence of effort, known as task commitment, is an important attribute for gifted students if they are to be productive (Renzulli, 1978). Renzulli defines task commitment as a more focused and/or developed form of motivation, the general energizing force that triggers behavioral responses. He views task commitment as "energy brought to bear on a particular problem (task) or specific performance area" (Renzulli, 1978, p. 182). He noted from the studies of eminent persons by Galton, Mackinnon, Roe, and Terman (cited in Renzulli) that "creative/productive persons are far more task oriented and involved in their work than are people in the general population" (Renzulli, 1978, p. 183). The highly creative architects studied by MacKinnon (cited by Renzulli) viewed themselves as enthusiastic, determined, and industrious.

Persistent practice of the varied creative thinking skills in CPS is seen by Torrance (1979) to be essential for attainment of potential. The experience of "satori," bursts of enlightenment which are the highest points of "expertness" (1979) is dependent upon commitment, discipline, and creativity. The Japanese, unlike Americans, value a concept of "creativity that requires perserverance, diligence, time, and hard work" (Torrance, 1979, p. 3). The Japanese have developed successful training techniques to develop these qualities and the "ability to delay gratification" (p. 9). Creative motivation, learning requisite creative skills, exercise of creative motivations (commitment), and a high level of creative ability are the essential components for creative behavior to be demonstrated (Torrance, 1979). Torrance recommended that creatively gifted, learning disabled youngsters should follow this advice: "Don't be afraid to 'fall in love with something' and pursue it with intensity and depth. You are motivated most to do the thing that you do best" (1982/1984, p. 20).

The relation of effort to perceptions of internal-external control and to creativity was discussed by Amabile (1983). She conducted an experiment in which children who were given a choice of materials to use in making collages were compared to children who were not given a choice. The children who were given a choice made collages that were judged to be more creative. She concluded that if people see their task engagement to be intrinsicially controlled, they will be more creative than if they feel their choices are extrinsically controlled.

Effort and Giftedness

Franks and Dolan (1982) reviewed theoretical and research literature describing educational impact of affective characteristics of gifted children. They focused on the importance of persistence, locus of control, field dependence, independence, and self-concept as the affective traits which are critical components for the emergence of gifted behavior. They noted that research on these relevant affective traits is limited by several factors. Many of the instruments to measure affective characteristics have inadequate validity and/or other measurement qualities such as unclear or unavailable scaling, norms, or developmental data. Moreover, basic research findings in the area of affect have in general not been incorporated into educational programming (Franks & Dolan, 1982). There is little consensus of operational definition of these three relevant characteristics. This last point is very well taken, for Franks and Dolan's definition of persistence is essentially what the present author has defined as effort (see definitions in Chapter I) which is distinguishable from persistence connoting the more restrictive meaning of task persistence.

Franks and Dolan (1982) cited studies which included persistence as a personality characteristic studied in comparisons between gifted and non-gifted children. Of various temperamental traits assessed by a self-report inventory, high persistence contributed significantly to the regression equation for reading achievement and for school adjustment in young gifted children (Burk, 1980). Dunn and Price (1980) found gifted students to consider themselves as more persistent than the non-gifted comparison group. Of the learning style characteristics studied by Dunn and Price, persistence ranked fifth in its ability to discriminate most between gifted and non-gifted students.

These findings, as well as the preceding analyses, suggest that although persistence (a willingness to do hard work) is associated with high intelligence, these appear to be independent traits. Franks and Dolan (1982) suggest that their review "supports the contention . . . that persistence may be necessary for the successful realization of high ability" (p. 173). They noted that although persistence in accomplishing goals is frequently considered to be an important aspect of giftedness, there were "relatively few studies which have considered this trait" (p. 174). In the present search of the literature, only a very few additional articles were located which considered effort in relation to creativity and/or gifted students.

Bennett (1984) studied affective characteristics in highly gifted elementary children who were classified as high achieving or as low achieving in reading or in math. Children completed two self-report ratings of their competence in reading and in math, and the Sydney Attribution Scale, a measure of self-reported attributions. The attributions of causal explanations for success or failure in academic situations were of ability, effort, or of causes external to the child.

Two experimental tasks were also administered to the children. Their preference for academic risk, that is, for their willingness to try difficult problems was assessed for vocabulary words and for math problems. Their willingness to expend effort was assessed by task persistence on decoding anagrams which were increasingly more difficult and then unsolvable. In addition to these measures, their classroom teachers rated the children on their persistence and their preference for difficult work as observed from class behaviors.

Classification as a low or high achieving student based on classroom reading performance bore little relationship to the children's performance on the experimental tasks for risk preference or persistence and did not show differences in causal attributions for success or failure. Classification by observed achievement in math bore a relationship to risk preference in math, that is, the high achievers in math showed greater willingness to try difficult math problems than did the low achievers.

Perception of competence in reading or math bore a closer relationship to performance on the experimental tasks than did their classroom performance as high or low achieving. That is, children who perceived themselves to be competent in math and/or in reading were more prone to attribute their success to their own effort and ability than did students who perceived themselves to be less competent than most of the others in their class. Students who felt themselves to be competent in math tended to seek the more difficult math tasks, and children who felt themselves to be competent in reading showed longer persistence after failure than those who felt less competent.

Contrary to the research findings with average learners, the attribution scale did not show differences in the gifted group in their causal attributions. Both high and low achievers tended to take responsibility for their positive experiences by considering their successes to be explained by their effort and ability. Both groups "took little responsibility for failure, weakly attributing it to effort, ability, and external factors . . ." (Bennett, 1984, p. 73).

Bennett's study provides an important validation of the role of effort in exhibiting gifted behavior. She suggested that the achievement in gifted students "might be enhanced by reinforcing greater feelings of personal control, more willingness to attempt difficult tasks, and a more positive academic self-concept" (Bennett, 1984, p. x).

A study by Sasfy (1976) of college students manipulated perceptions of causality for their participation in an experimental creativity task. The subjects whose perceptions of causality were influenced toward intrinsic sources of causality were found to have greater motivation and enjoyment in the task than those subjects who were influenced toward extrinsic or social sources of causality (Sasfy, 1976).

Effort was a variable considered in a study by McKenna (1981). Scores on the IAR were utilized as the measure of perception of responsibility for academic achievement of gifted, bright, and average students. Effort was defined by a separate report card grade for effort in reading as distinct from a grade for achievement in reading. McKenna discussed the relationship between effort and locus of control in terms of attribution of causal explanation of successes or of failures.

> Pupils who tend to attribute success to ability and effort (internal causes) experience pride for their successful performance and therefore tend to approach achievement tasks. In contrast, pupils who tend to attribute success to external causes experience less pride for their success and therefore tend to avoid achievement tasks (McKenna, 1981, p. 9).

Moreover, of the internal causes these can be further classified as stable or not readily modifiable, such as ability, or as unstable, such as effort. Therefore, those who attribute failure experiences to insufficient effort, an unstable internal cause, would be likely to persist when faced with failure at a task because they believe they can change the outcome by modifying their own effort. Those

> . . . who believe that the outcome is greatly determined by effort {will} perform tasks with greater intensity than those who believe that the outcome is determined by external causes. The belief in effort makes the pupil try harder, because effort is believed to be an unstable-internal cause (McKenna, 1981, p. 10).

McKenna's finding that locus of control was related to achievement but not to intelligence is consistent with the work of Kanoy et al. (1980). No other comparison with the IAR measures were significant other than a positive correlation between I+, I-, and I_t with the grade for effort in reading. There was some tendency for the I- scores, responsibility for negative events or failure situations, to be lower for the gifted and superior students than for those of other intellectual categories. This trend and the variability of the finding is consistent with the results of Bennett's (1984) study. Items which may typically be responded to by the average child as taking responsibility for negative events may be less relevant to the experiences of the gifted child.

McKenna (1981) consistently found the effort grade to be judged as equal or higher than the achievement grade in reading for all intellectual categories at Grade 3, but that effort, as graded by their teachers, was judged to be lower than achievement for those Grade 5 students who were identified as mentally gifted or superior. She interpreted these findings to be consistent with the work of other authors in which teachers' behavior, when students fail easy tasks, tended to be more punishing toward students perceived by their teachers as competent or high in ability than toward children who are low in ability.

Self-Efficacy and Creative Behavior

The construct of self-efficacy describes a cognitive mediator which is the person's belief in his or her ability to be competent at a behavior in a specific situation. The perception of self-efficacy is specific to given situations, and will affect decisions to initiate the behavior, effort and persistence in the behavior. Starko (1988) developed a set of instruments to assess self-efficacy and creative productivity in gifted students. The students were identified gifted middle school students, some of whom had been in gifted programs based on the revolving door identification model developed from the enrichment triad model of Renzulli (1977). The students who received the gifted services were compared to those who had not been served: by their participation in creative productivity outside of school; by a number of major original products completed in school (Type III creative products; Renzulli, 1977); and by scores on the self-efficacy scale.

Analysis of hierarchical multiple regression equations revealed that group membership and number of creative products produced in school were significant predictors of creative productivity outside of the school setting. Self-efficacy toward creative productivity correlated significantly with creative productivity in school.

Additional validation of the construct of self-efficacy in regard to creative productivity was provided by the work of Burns (1988) and Schack (1987). Burns found self-efficacy to be one of eight predictor variables which significantly separated students who did or did not initiate creative products. Schack, likewise, found self-efficacy to be a significant predictor of creative behavior by predicting initiation of original (Type III) projects for children in Grades 4-8. It is possible that the methodology of this research could be subject to the criticisms of creativity research using divergent thinking measures leveled by Mansfield et al. (1978); that is, there may be too much similarity between the training variables and the criterion variable. Nevertheless, the construct of self-efficacy toward creative productivity contributes several important aspects to research based in creativity. It defines effort and persistence as separate, related aspects of the construct, and affirms the thesis presented earlier in the study (pp. 11) that affective measures need to be specific to the content of the task.

Some Operational Definitions of Effort

Children's self-perceptions of their ability, effort, and conduct were evaluated to determine whether these were distinguishable characteristics (Blumenfield, Pintrich & Hamilton, 1986). Children responded to open-ended questions, one of which asked, "How do you know when someone is a hard worker" (Blumenfield et al., 1986, p. 98). Responses to this question centered primarily on the behaviors of "finished work" and "always working," and to a lesser extent, "stays out of trouble," "works fast," and lastly, "works hard" (Blumenfield et al., 1986, p. 98). Their results suggested existence of strong intercorrelations among the three characteristics. A regression analysis showed that when effort and conduct were entered to predict ability, "effort remained a significant predictor (beta=.31, $p < .001$) but conduct dropped to insignificance" (Blumenfield et al., 1986, p. 101). These findings provide experimental validation that children's perceptions of effort are, as suggested by Renzulli (1978), closely tied to the construct of ability and gifted behavior.

Friedman, Friedman and Van Dyke (1984) conducted a study to predict gifted leadership performance ratings from self, peer, and teacher nominations. They equated leadership giftedness with Renzulli's (1978) construct of giftedness as comprised of three criteria: above-average ability, creativity, and task commitment. The fourth and fifth grade students were nominated for leadership committees to help plan a special electives program. The 28 student committee members were given tasks that were rated on three criteria: group leadership ability, creativity, and task commitment. Of particular interest to this study is the method by Friedman et al. of arriving at a definition of "leadership ability" and of task commitment.

Leadership ability was determined by peer ratings of the other six committee members in response to these Likert-type questions: "How much they helped the group to get the job done well," and "How much they helped the group to feel good about meeting together" (Friedman et al., 1984, p. 92). The ratings derived from the responses to these two questions were combined into a 10-point leadership ability score. Task commitment was determined by the number of hours (0-10 hours) that the children stated they could be counted upon to do work for the special electives classes. These scores were added to the creativity score (a list of new topics suggested by the committee members for sessions of the special electives classes). These topics were scored for fluency, flexibility, and for originality.

Friedman et al. (1984) were primarily interested in assessing the effectiveness of the three different sources of nomination as predictors of the above operationally defined score for leadership giftedness. They found that students who were nominated by all three sources, peer, teacher, and self, showed the highest ratings of leadership giftedness and that students who were nominated by two sources obtained a higher leadership giftedness score than those nominated by any single source, especially if they were nominated by peer and self or by teacher and self. There was little difference among the three single nomination sources as effective predictors of leadership giftedness.

The present author considers that Friedman et al. (1984) have validated a way to offer an operational definition of task commitment and of leadership ability. Their methodology showed that use of peer ratings of others' contributions to the group's effectiveness and of self-ratings of willingness to commit time to a task provided an effective operational definition of their constructs.

As can be seen from the preceding review, effort is considered to be a critical component in the enhancement of creative production (Osborn, 1957; Torrance, 1979). It also interacts with creative abilities and mental ability to enhance gifted behavior (Franks & Dolan, 1982; Renzulli, 1978). A few research studies have verified differences in effort associated with high and low achievement in gifted students (Bennett, 1984; Burk, 1980; Franks & Dolan, 1982; McKenna, 1981). Other authors state that many gifted students frequently put small amounts of effort into their work while maintaining successful grades and consequently fail to learn good study habits, and/or a healthy attitude toward work and competition (Clark, 1983; Rimm, 1986).

Effort and Odyssey of the Mind Teams

Most of the gifted students who participated in what was then called Olympics of the Mind described it as the most challenging of the activities in the gifted program and the one in which they experienced the "greatest feeling of accomplishment, success and satisfaction" (Miller, 1983, p. 50). A small number, however, indicated that it was the activity in which they felt the least success and/or least challenge. Miller's findings suggest that an evaluation or research study of effectiveness of an OM program might find it productive to separate students in OM based on some variable similar to satisfaction, achievement, and/or challenge. Because amount of effort people put into a task is related to their perception of the task and to their feelings of enjoyment (Sasfy, 1976), it is possible that differences in effort on an OM team could be a variable that could help to explain Miller's findings.

Effort was operationally defined in relation to grade point average (McKenna, 1981). Effort has also been operationally defined as willingness to devote time to the work for a special program (Friedman et al., 1984). However, there has not yet been a study in which effort is operationally defined and used to compare high effort and low effort children for differences in their behavior on a creative task. The present author suggests that, based upon the preceding review, differences in effort will occur among members of a given team and across teams. Moreover, such differences in effort ought to account for differing patterns of growth in cognitive and affective measures which might describe the creative child. If, indeed, increased persistence of effort is involved in effective work on an OM team, those who put extra effort into their OM teamwork would be expected to show greater gains in creativity than those who do not exert the effort.

Experience in a Program

Because of the novelty of the creativity training program one might predict that a greater effect would be found in a group of children who are new in a program than in those who have had previous experience. Differing amounts of prior experience with creativity training would be likely to have differing effects on measured creativity.

A study by Jones (1982/1983) examined the effects of experience in an enrichment program on elementary mentally gifted students in their first, second, or third year in the program. Six ANOVAs were used to analyze the data presented as mean gain scores from pre- to posttesting after nine weeks of attendance in a gifted seminar program. Significant differences were found in gain scores across all subjects in all six criterion variables, with effects on total IQ, TTCT verbal, TTCT figural, Piers-Harris Children's Self-Concept scale, and Metropolitan Achievement Test (MAT) scores. The effect of differences in length of time in the program were observed on three of the six scales, ANOVA of MAT tests and were most pronounced on the creativity tasks. Post hoc tests of the interaction of creativity tests located the greatest gains in the first year, then the second, with little gain in figural scores in the third year and none in the verbal scores. Jones indicated that levels of experience in the seminar program made a difference and a cumulative effect seemed evident on scores of creative thinking tasks and particularly on figural scores. The verbal task performance of students in the third year suggested that changes were slowing or perhaps might even have stabilized at a high level that is appropriate for gifted and talented children (Jones, 1982/1983).

It is interesting to note that the greatest change in creativity scores occurred during the nine-week pre- to posttesting interval in decreasing order in relation to student experience in the program. Two possible explanations, not considered by Jones (1982/1983), arise from the review of the literature in this study. Jones' samples contained a confounding of the variables of age and grade (fourth to sixth) with the experience factor. Perhaps the phenomenon of greatest growth in the first year may be partially explained by provision of a program which offset the "fourth grade slump" (Williams, 1976; Torrance, 1972a). An alternative, or perhaps confounding, effect could be produced by similarity of testing to the treatment process, an artifact of studies utilizing the TTCT that was noted by Mansfield et al. (1978). "Success of a creativity training program may be due only to an increase in persistence or to the knowledge that original and clever responses are wanted" (Mansfield et al., 1978, p. 518).

The degree of experience in a gifted program appears, from Jones' (1982/1983) study, to make a difference in the amount of change to be noted due to the treatment, with naive subjects experiencing a greater effect. Parnes (1961), however, found that subjects who were experienced with CPS training techniques produced more ideas and higher quality ideas in solution to a problem than novice subjects. Although both groups were given the same instruction to defer judgment, to interact freely, and to try to think of as many alternative solutions as possible, those who were experienced in CPS outperformed those who were new to CPS. An extended CPS training progrm might, however, when evaluated for change over time, have a greater impact on those who were initially new to the CPS techniques.

Summary

In this chapter an argument was developed for the need for research on the effectiveness of team CPS with gifted children in its effects on affective as well as creative behaviors. The review of the literature surveyed theoretical issues in the field of creativity which pertain to definition and measurement of creativity, of relationships between creativity and the gifted, and, more specifically, CPS. Theoretical issues related to definition, specificity, and measurement of self-concept and locus of control were also surveyed. The literature was examined which related to CPS, affective characteristics as effects, of CPS and the gifted.

The nature of social aspects in groups which might be pertinent to the team CPS process was included as a possibly relevant factor to this study. Issues pertaining to effects on creativity of working in a group, collaboration, competition, and perceptions of evaluation were discussed.

Studies defining task commitment and effort were reviewed in the context of their relationship to creative behavior and/or to gifted children.

CHAPTER III

METHODOLOGY

Introduction

This study utilized multivariate mixed-model analyses of two separate multivariate constructs: creativity and affect. The independent variable was effort in an OM CPS team. Supplemental analyses were also performed.

Subjects

All subjects were mentally gifted children with individually assessed IQ scores at or above the 97th percentile. Each student was enrolled in the pull-out enrichment program in one of the ten elementary schools of the district under study. The sample consisted of 143 gifted students in grades 2-5 whose parents gave permission to be tested for the study. Permission was not granted for 10 students who were not tested in this study. Complete data could not be obtained from an additional 16 subjects resulting in an attrition of 11%. Complete data was obtained for all tests of either the creative or the affective test battery for 117 children; with 116 of these 117 children comprising the creative sample and 111 of the 117 comprising the affective sample. There were unequal numbers in the OM groups. The affective sample consisted of the following: not in OM (non-OM)=49, OM team members who were low in student-perceived effort (OM-lo)=29, and OM team members who were high in effort (OM-hi)=33. The creative sample consisted of the following distribution: non-OM=51, OM-lo=32, and OM-hi=33. Subjects were unequally distributed among groups with respect to several additional uncontrolled variables: grade level, sex, months of experience in the enrichment program, years of prior experience in OM, and home schools in which the three itinerant enrichment teachers held their classes. The distribution of subjects in the cells of the independent variable of this study is described more fully in Appendix F. The distribution of subjects with respect to the variables of effort in OM and experience in the enrichment program is described in Table F-1, and with respect to prior experience in OM in Table F-2, Appendix F. The distribution of subjects with respect to effort in OM and grade level and sex is given in Table F-3, and with respect to teachers, effort in OM, and home schools in which their classes were located is in Table F-4, Appendix F. All of the subjects for whom complete data was available were used in order to obtain adequate power. Because of findings from a pilot study (see Footnote 1), it was assumed that the OM x Time interaction ought to be of a moderate effect size. Therefore, with $\underline{n}=29$ in the smallest cell, power should be at .58 with $\alpha=.05$ for interaction effects in this study (Cohen, 1977). Power, that portion of the distribution of the research hypothesis which represents difference(s) that are real, (Tabachnick & Fidell, 1983) is further defined in Appendix B.

Sources of sampling bias which could affect generalizability of the study include factors unique to the socio-economic level of the community and local methods of identifying intellectually gifted children. The district serves a city of moderate to high socio-economic status. Students are qualified for the enrichment program by attaining a score on an individual test of intelligence at or above the 97th percentile. Evidence of giftedness in academic functioning or creativity may be considered by the referring teacher, but it is the policy of this

Note 1: A pilot study of creativity scores obtained a moderate effect size (.23) for the OM x Time interaction and a very large effect size for time (> 1.0).

district that no criterion other than IQ is considered for placement into the program (see Appendix C).

Instruments

Instruments assessing the dependent variable constructs of creativity and affect were administered as pre- and posttests for the repeated measures. The instrument used to define the independent variable, effort in OM, was the Rating Scale of Work Done by the Team administered during the first week following the local competition.

Permission to collect data on instruments which were not part of normal enrichment program activities was obtained from appropriate administrators in the school system. A letter requesting parental permission (see Appendix G) for their child to participate in the study was mailed to all enrichment program parents on September 4, 1985. The potential sample size consisted of 143 students as of September, 1985. A response period of two weeks was requested to return the permission form in an enclosed, stamped envelope. After the indicated return date, a telephone follow-up call was made to the parents of 46 children who had not yet responded by October 1. Of these, verbal permission was given for 40 children, written permission for 41 of the children was received within six weeks. Another five forms trickled in during the school year to obtain a final count of written responses for 140 children, a 98% return rate. The overall rate of approval given by the parents for their child to participate in the study was 93% (see Table F-5, Appendix F).

The letter of September 4 did not request permission for individual testing on the TTCT for second and third grade children. Parents of the 47 primary grade children were phoned during the first two weeks of October. All gave verbal permission for individual testing, and written confirmation (see letter in Appendix G) was received by November 14 for 46 of the 47 children.

Effort in Odyssey of the Mind

Measurement of Effort

Personal observation of children who have participated on OM teams in past years indicated that most of the children were highly involved, exerted a great deal of effort on their team, and enjoyed and benefited from the experience. Many of the children seemed to be more creative, more confident in their abilities, and more self-directed. These observations are consistent with OM research findings cited in Chapter II (Harrington, 1984; Miller, 1983). Miller, however, found that a small minority of the students experienced less success in OM than in any other of their activities. Cohen (1987) found differences within teams, among teams, and in coaches in creativity, self-concept, and other attributes. Likewise, a few of this author's students who had been in OM continued to find difficulty with offering their ideas to a group and/or difficulty with refraining from criticism of others during group brainstorming. Some of these were students who indicated that they did not wish to be on an OM team again. It therefore occurred to this author that perhaps important differences in effort exerted by team members might be contributing to these observed differences. Individual differences in amount of effort invested in the team's work would help to explain some of the conflicting findings of increase in creativity associated with working in a group, as discussed in Chapter II (pp. 82-86) (Amabile, 1983; Bolen & Torrance, 1978; Roweton, 1982).

A search of the literature prior to designing this study (see Chapter II, p. 96) found no appropriate precedent for assigning students to high and low effort groups for OM. The students on OM teams did not do their work with their enrichment teacher because, in this district, OM is offered as an extracurricular option (see procedures and Appendix D). Therefore, the three teachers were not in a position to make an informed observation of amount of effort expended by the children in their OM work. The 83 children in OM were on 17 different teams where each team was facilitated by different volunteer parent coach(es). (The number of students in OM in the study was 65; most of the 18 children in OM who were not in the study were children who enrolled in enrichment after mid-September.) It was thus not possible to obtain a consistent base of adult observation to evaluate the children's effort on the 17 different teams.

It was, therefore, decided to use the children's rating of their own and their peers' effort. The children in this program are exposed to other experiences in self and peer rating, such as rating student presentations of independent research projects. The study by Friedman et al. (1984), discussed in Chapter II (pp. 94-95), found that self and peer ratings of leadership work demonstrated in a problem solving team, created a leadership ability score which was a valid component of an overall leadership giftedness score comprised of leadership ability, creativity, and task commitment. Peer rating is a valid method of psychological assessment. Peer ratings among adolescents obtain correlations from .45 to .70 between reputations and corresponding behaviors (Cronbach, 1960).

Because it was necessary to obtain a score on each team member, a ranking method rather than a rating scale was devised. Requiring the children to rank themselves and each other member of their team on each of three questions provided a method which limited facade or social bias effects (Cronbach, 1960). However, some estimate of overall amount of work performed by the team as a whole was needed. Without such a weighting factor, the resultant scores could be higher for a child perceived by her team to be the hardest worker when the team, as a whole, did not produce much work than the score of a child who worked moderately hard on a team where the entire team worked hard. Therefore, the first question, "How much work was completed by your team?" was included to provide an adjustment weight to self and peer ranking scores.

The Rating Scale of Work Done by the Team is a self-report and peer ranking scale (see Appendix H). A score of 0, or no effort in OM, was assigned to the non-OM group. OM-lo and OM-hi groups were determined by below or above median scores for each child's total rating on the three questions which ranked an individual member's effort, after the scores were adjusted by the mean assessment of work performed by the team as a whole.

The instrument consisted of four questions. The first assessed work by the team as a whole and three additional questions rated each member of the team. The three questions asked the children to rank themselves and each member of the team on time spent, completion of work and ideas contributed to the team. For each of the three questions, children were asked to place the names of each member of their team including themselves in order. The names placed on the list were assigned a weight from a low of 1 to a high of 3 for each question. These ranked scores were thereby adjusted to account for differences in team size (see calculation of the effort score, Appendix H). To account for differences in effort among the teams, the mean weight of work done by the team, as perceived by each member of the team, was added to each member's score. The range of possible scores was from 4 to 14. The obtained range of scores was 7.2 to 13.3.

Inter-rater reliability coefficients for the three to six respondents per team were computed for the effort score which resulted from the three self/peer ranking questions added to the child's perception of work completed by the team. Within team Spearman rho correlations compared each team member's effort score to the composite ratings of the remaining members of the team. These within team inter-rater reliability coefficients yielded a range of -.50 to .95 and median $r=.57$ for the 17 teams, with an overall correlation of .99, $p < .01$, for the

81 students on the teams, thus providing satisfactory evidence of reliability for the effort scale.

As it was not possible to obtain teacher ratings of the children's efforts in their work on their OM teams, it was necessary to use some other methods of estimating the validity of the effort scores. Two independent measures were constructed which could provide some estimate of the children's effort on a CPS team. Because the enrichment teachers could observe the children's effort and teamwork skills when they were in class on Future Problem Solving teams, a teacher rating scale was devised to yield scores for each child of their effectiveness in their within class teams. It must be noted that Teacher Rating Scale of Team Effectiveness (see Appendix I) rated the children's work in a situation that was only an approximation of their OM work. Work on OM and Future Problem Solving teams is similar because both situations encourage considerable interaction, brainstorming, and consensual decision-making in a situation where a considerable amount of hard work and shared effort are required. The coaches in OM and Future Problem Solving serve in a facilitative role to help the children in their process. However, it is quite possible in this setting that the teacher coaches may be more adept than many of the volunteer parent coaches in helping the teams to learn to defer judgment, and therefore to learn to resist criticizing their own and others' ideas, to involve reticent members of the team, and to help the teams stay on target. Evaluations by parent coaches from the past several years often indicated that one of their greatest needs for further training was for techniques in team discipline.

It must also be noted that team composition was different for the OM teams than the Future Problem Solving teams. The Future Problem Solving teams consisted of 2-5 children in the same enrichment class. Moreover, the teams were likely to be comprised of different children for the two sessions of Future

Problem Solving work. Almost none of the OM teams were composed of membership identical to either of the Future Problem Solving teams. The OM teams were to some extent self-selected with general guidance from their teachers based on interest in OM, in the particular problem, and in working with the other children who were considering that team. In order to fill out the teams with enough members who were all interested in working on a given OM problem, it was necessary to find at least one member from a different school for five of the 17 teams. At least four additional teams had one or more members who attended a different enrichment class than the others on the team. It must also be noted that it is possible that OM teams may be somewhat more likely than Future Problem Solving teams to be composed of children from different grade levels. Most of the OM teams consisted of children in Grades 4 and 5 or Grades 2 and 3, but two upper grade teams had a third grader as a member.

As a concurrent measure to assess the validity of the effort scores as derived from the Rating Scale of Work Done by the Team, the teachers evaluated the children's effectiveness and effort in the gifted classroom team CPS activities on Future Problem Solving teams. The Teacher Rating Scale of Team Effectiveness is a 16-item Likert-like scale (see Appendix I). The three teachers evaluated each of their own students within two weeks after the teams completed the work on each of the two Future Problem Solving Program problems. The resultant scores for the two ratings for each child by their own teacher yielded a mean score of teacher rating. A correlation was calculated between the peer/ self-ranking measure, Rating Scale of Work Done by the Team, with the teacher rating for each child who was on an OM team, Spearman-Brown, r_s=.20, non-significant.

A second concurrent measure of validity was the Student Evaluation of OM (see Appendix J). This self-rating Likert-like scale contained some questions to parallel the Teacher Rating Scale of Team Effectiveness and the Rating Scale of Work Done by the Team. The scores on this measure were available for only 29 of the 83 elementary children on an OM team. The scores of those fourth or fifth grade students that were available yielded a Spearman rho correlation with their self-rating score on the Rating Scale of Work Done by the Team of $r=.45$, $p < .05$. This latter comparison provides evidence of concurrent validity for the effort scale.

Measurement of the Creativity Construct

Based on the evidence of the effectiveness of CPS training to increase TTCT scores, this study will use both the verbal and figural batteries of the TTCT. From the review of the literature (pp. 40-47) it seems that a moderate to large treatment effect would be expected on verbal creativity scores of students who received OM training. Verbal creativity would be expected to increase because all students also received CPS training in their work on Future Problem Solving units. A variety of training experiences in figural creativity as well as in figural problem solving are also incorporated into the ongoing enrichment program; therefore, figural scores would also be expected to increase.

The multivariate construct, creativity, was measured by three instruments: the Torrance Tests of Creative Thinking (TTCT) figural battery (Torrance & Ball, 1984) and verbal battery (Torrance, 1974a) and the Similes Test (Schaefer, 1971). These yielded six figural and four verbal measures of creative thinking to form the multivariate creative construct.

Torrance Tests of Creative Thinking

The Torrance Tests of Creative Thinking were first published in 1966 and are now available, under copyright, from Scholastic Testing Service. Torrance made deliberate attempts to develop tasks designed to model the creative process and to elicit a variety of creative behaviors. According to Torrance (1974b), the tasks selected for the figural and verbal batteries show only a low correlation in factor analyses. Because of this low correlation, the final tasks, then, are each considered to involve a different kind of behavior and contribute a unique quality to the battery (Torrance, 1974b).

The TTCT is appropriate for use for kindergarten through graduate school. Clear directions for administration and recommendations to standardize motivational conditions are provided in the manuals. Working time for students is 30 minutes for the figural and 45 minutes for the verbal. The manner of presentation is designed to elicit curiosity and maintain a game-like atmosphere. The figural tests are group administered for all age levels. The verbal tests are group administered at fourth grade and older but must be individually administered at kindergarten through third grade.

Verbal Tests: Standard Scoring. Reliability data for the verbal TTCT scoring (Torrance, 1974b) show interscorer reliabiities between untrained classroom teachers who studied the scoring guide and experienced scorers to be high, ranging from .88 to .99 for verbal booklets, and .66 to .99 for figural booklets scored according to the original standard scoring procedures. Equivalency test-retest correlations using an alternate forms design (AB, BA) for all verbal and figural batteries ranged from .50 to .87 in a sample of 54 fifth graders and .71 to .93 in a more diverse group of 118 4th-6th grade students. Other test-retest reliability coefficients from studies using a variety of abbreviated batteries were generally above .60 (Torrance, 1974b). Chase (1985) and Treffinger (1985) considered the retest reliability reasonably adequate for evaluating changes in characteristics for group or research uses. Reliability and validity issues regarding the TTCT are discussed in greater depth in Appendix E.

The procedures for this study were originally designed to use a proposed streamlined scoring for the verbal tests (Torrance & Ball, in press), as well as the streamlined figural (Torrance & Ball, 1984). As recommended by the forthcoming manual for the administration and scoring for streamlined verbal procedures (Torrance & Ball, in press), Activity 6 "Unusual Questions" was not administered in this study. It was later determined that the streamlined verbal scoring procedures were not quite ready for use. The verbal responses were, therefore, scored according to the standard scoring procedures for the remaining six of the seven verbal activities which were administered.

Verbal test booklets were scored by the scoring service of Scholastic Testing Service which used the standard scoring procedures. An inter-rater reliability was performed by the scoring service by scoring each of the Verbal A booklets independently by two different scorers. The inter-rater reliability obtained by experienced scorers for the sample in the present study was: verbal fluency=.95, verbal flexibility=.86, and verbal originality=.86. These analyses were based upon standard scores (E. J. Stehlin, June 18, 1987, personal communication, see Appendix A).

Figural Tests: Streamlined Scoring. Previous data on scoring reliability, implied to mean inter-scorer reliabilities, was reported for the figural battery (Torrance & Ball, 1984). Reliability coefficients obtained from two studies at the elementary level, grade 2 and grade 5, for the five norm-referenced scoring variables--fluency, originality, abstractness of titles (titles), elaboration, and resistance to premature closure (closure)--were all at least .90 for grade 2 and .98 for grade 5. The 12 criterion-referenced variables are: emotional expressiveness, storytelling articulateness, movement or action, expressiveness of titles, synthesis of figures, synthesis of lines, unusual visualization, internal visualization, humor, richness of imagery, colorfulness of imagery, and fantasy. Most of the creative strengths obtained reliability coefficients over .85 (see discussion in Appendix E). Product-moment correlations were calculated for the same scorer using the original scoring procedures, and six months later, for the streamlined procedures, and resulted in inter-rater reliability for fluency, originality, and elaboration scores of .92 to .94.

For this study all figural booklets were scored by the author using the streamlined scoring procedures (Torrance & Ball, 1984). The inter-rater reliability was established by independent scoring of a portion of the booklets with three outside experienced scorers. A sample of 40 of the figural booklets, 20 pretests and 20 posttests, was scored by the professional scoring service of Scholastic Testing Service. The 20 pretest booklets scored by the scoring service were used along with the manual and workbook (Ball & Torrance, 1984; Torrance & Ball, 1984) in a self-training process by this author. After scoring all the Figural A and Figural B, the inter-rater reliability was calculated on the 20 Figural B scored by this researcher and previously scored by the scoring service. The correlations of the standard scores for the five norm referenced measures and the raw score for total creative strengths ranged from $r = .63$ for elaboration to .99 for fluency (see Appendix F, Table F-6). The reliability coefficients for closure, elaboration, titles, and total of creative strengths were considerably below those reported by Torrance and Ball (1984). The 20 booklets were, therefore, scored a third time by a graduate student at the University of Georgia. The raw score means, standard deviations, and reliability coefficients are reported in Table F-6 and standard score comparisons are reported in Table F-7, Appendix F. These scores indicated the likelihood of an error in scoring originality by the Georgia rater and possible differences in interpretation of scoring criteria for four other measures--titles, closure, elaboration, and total of creative strengths--that suggested a close correlation of scores between this rater and the Georgia rater which were both discrepant from the ratings obtained from the scoring service.

Because the streamlined manual (Torrance & Ball, 1984) was under revision and discrepancies were obtained in the inter-rater correlations, STS offered to rescore the booklets by their senior scorer. They suggested that the source of the discrepancies could perhaps be interpretation of the manual in areas where the manual was in need of clarification for the current revision (J. D. Kauffman, February, 1987, personal communication, see Appendix A). This final inter-rater correlation was calculated by the scoring service on the raw scores rather than the standard scores (see Appendix E). The inter-rater reliability coefficients of raw scores obtained by this author and the Scholastic Testing Service senior scorer were: fluency=.99, originality=.96, titles=.82, elaboration=.76, closure=.89, and total of creative strengths=.66. For further discussion of inter-rater reliability issues on the TTCT see Appendix E. The means and standard deviations for these comparisons, reported in Table F-8, Appendix F indicate a consistent overestimation of the elaboration criteria by one rater, as discussed in Appendix E. The final set of comparisons of ratings between this author and the senior scorer of the Scholastic Testing Service, despite the overestimation of the elaboration measure, show evidence of satisfactory inter-rater reliability for this study.

A pretest posttest comparison of alternate forms reliability was calculated for the verbal and figural measures and the results are reported in Table F-9, Appendix F. These correlations for the verbal measures are: fluency=.49, flexibility=.51, originality=.45, and mean verbal=.53; for the figural measures are: fluency=.37, originality=.45, titles=.32, elaboration=.42, total of creative strengths=.40, and Creativity Index=.45, all at $p < .01$, and closure=.18. These reliability coefficients, however, should be considered as underestimations of the alternate form reliability for the TTCT for several reasons, including the following: 1) Significant treatment effects were obtained for the creativity measures which differed in relation to group membership according to the independent variables of this study. 2) The reliability coefficients for this study were calculated on standard score conversions for the verbal measures and for the five norm-referenced measures of the figural test. Use of standard scores, when available, are usually preferred to use of raw scores because the standard scores are the most powerful scores for use in statistical comparisons provided that the data are not markedly skewed (Isaac & Michael, 1981). It was later learned that TTCT reliability studies are normally performed on raw scores instead of the newer standard scores, in part because the conversion procedures introduce an additional source of unreliability to some of the measures, e.g., closure and titles (see further discussion in Appendix E). However, because the criterion validity of the TTCT is high (see Chapter II and Appendix E), low obtained reliability is not considered to be a problem (Isaac & Michael, 1981).

Verbal Measures: Similes

The Similes Test (Similes) by Charles E. Schaefer (1971) is a copyrighted measure from Research Psychologists Press. Its purpose is to identify literary talent in children and young adults. There are two forms of the instrument consisting of ten incomplete simile stems such as "The cat's fur was smooth as __________." Subjects are asked to give three endings to each similes stem in the 15 minute time limit. The example given in the instruction implies that responses may contain more than one word. Reasonably extensive norms are available for students in Grades 4-6; and high school, college and adult norms are also provided.

The test may be given by persons familiar with standardized testing procedures. The guides for scoring Similes are intended primarily for use by teachers and others who need not be expert in testing or in evaluating poetry (Schaefer, 1971). A rating scale guide is provided in the manual with clear guidelines. It is possible to establish interjudge reliability of .93 for Form I and

.98 for Form II for experienced scorers (Schaefer, 1971). The scores for scorable responses range from 1 for a response that is merely descriptive and ordinary to 5 for a response that is not only unusual but highly appropriate and captures the essence of the simile stem.

Odd-even reliability coefficients (using the Spearman Brown prophecy formula) in the .80s are reported. There is not yet adequate data to support a test-retest reliability with use of the alternate forms. A study utilizing the alternate forms after two weeks with bright fifth graders yielded a stability correlation of only .60 (Schaefer, 1971).

Reviewers in the Eighth Mental Measurements Yearbook are in agreement that the test has adequate reliability, inadequate evidence of a good correspondence of the alternate forms, and is in some need of further validity studies (Clark, 1978; Frederiksen, 1978). Therefore, Form I was used for pre- and posttest measures. An inter-rater reliability was conducted for a random sample of 20 booklets. The raters were the present author and three students in a graduate class in creativity. Pearson correlation coefficients of the present author with each of the three other raters were: .92, .92, and .94. Reliability coefficients of stability were obtained by correlating the children's performance on their pre- and posttests with and intervening time interval of six months. The obtained test-retest reliability coefficient of .53 should be considered to be an underestimation of the stability coefficient for this test (see Table F-9, Appendix F). Moreover, a treatment effect was obtained for differences in Similes posttest scores associated with group membership, which contributed to a lowering of relationship between the pre- and posttests (see Appendix E for further discussion of test-retest reliability procedures).

Measures of the Affective Construct

Three measures of conceptually different affective scales were utilized:

general self-concept, Sears Inventory (Sears); creative self-concept, How Many Ideas? (Ideas); and the two scales of the Intellectual Responsibility Questionnaire (I+ and I-). The measure of creative self-concept was expected to bear the closest relationship to the creative thinking measures and therefore to the hypothesized changes in creativity and feeling good about oneself as a creative, productive person. The measure of locus of control was expected to reflect confidence in one's own ability rather than that of other people to bring about changes. The locus of control measure was expected to bear a closer relationship to the amount of effort exerted in student and team-directed CPS and was therefore expected to reflect hypothesized changes in the student's sense of internalized power. The general self-concept measure, as a more global characteristic, was expected to be a more stable measure.

Creative Self-Concept

The creative self-concept scale, How Many Ideas? (Ideas), is a 10-item Likert-like scale developed to assess a child's view of the self as a producer of creative, productive, and original ideas and/or products (See Appendix K). It was adapted from a similar self-report scale used in a study of creativity and general and creative self-concept in college students (Wright et al., 1975).

The Ideas scale (Fishkin, 1987a, 1987b) was developed because no other scale assessing the specific self-concept in relation to creativity was found. The review of the literature indicated a need to assess the domain of self-concept specific to creativity. The 10-item scale used in the Wright et al. (1975) study measured aspects of the self more highly related to creative behavior than the general concept of self. The scale was rewritten by the present author to clarify items and lower the reading level. Analyses of the reading level yielded grade level estimates of 7-8 for the Dale Chall and 5 for the Flesch readability formulas. Of the two formulas the Flesch is considered to give a more reliable estimate. Adjustment factors of -.285 to be subtracted from the Dale Chall scores have been suggested (Guidry & Knight, 1976; Rush, 1985) therefore indicating that the reading level of this instrument is somewhere between 4.7 and 7, within the reading level for the gifted children in this study. Responses were weighted from 1 to 5, with "5" representing "often" descriptive of the respondent for one half of the items and "seldom" for the other half. Reliability tests of Ideas yielded an internal consistency coefficient of $\alpha = .70$ for 115 of the pretests of the children in the study and $\alpha = .85$ in a sample of 21 graduate students.

As discussed above and in Appendix E, a test-retest reliability coefficient was computed for this measure as well as the other five instruments in the study, which were administered as pre- and posttests. For reasons discussed above and in Appendix E, the obtained stability coefficient of .44 (reported in Table F-9, Appendix F) should be considered as an underestimation of the stability coefficient if it were to be obtained over a shorter time interval without intervening significant treatment, and if it were to be obtained from a sample likely to have greater variability of scores in creative self-concept. Ideas, then, represents a measure of sufficient reliability that is appropriate for use by elementary-aged children.

The content validity was established by having the scale examined by persons knowledgeable about self-concept and creativity. Concurrent validity of the parent creative self-concept scale was established by its relationship to scores of the Tennessee Self-Concept Scale and the scores of the TTCT (Wright et al., 1975). Criterion and construct validity of Ideas will be assessed by the results of this study and will be discussed in Chapter V.

Locus of Control - Intellectual Achievement Reponsibility

Locus of control is a personality characteristic different from general self-concept or self-esteem. This construct describes the degree to which a person accepts or feels appropriate responsibility for his own actions. It has been known to predict other behaviors such as task persistence, effort in academic tasks, performance on achievement tests, and creativity (Crandall, 1978). Persons with a higher internal locus of control, then, would be expected to show greater commitment to participation in their freely chosen activities. However, one may find that in team selection sometimes children may feel they were drafted and the process, therefore, may not have been perceived as their freely chosen activity.

The Intellectual Achievement Responsibility Questionnaire (IAR), which measures locus of control, was first reported in Crandall, Katkovsky and Crandall (1965). The original form of the IAR is a 34-item forced choice scale suitable for Grades 3-12 (See Appendix L). The test items consist of a routine positive or negative achievement situation described as an item stem. The stem completion alternatives imply either that the event was caused by the child or by someone else in the child's life. The test yields three scores: I+=internal responsibility for positive events, I-=internal responsibility, or acceptance of blame, for negative events, and I_t=total IAR. The simplified form of 32 items is recommended for children below third grade (V. C. Crandall, personal communication, July 7, 1985). A short form is composed of only 14 of the items which are more specifically related to effort (V. C. Crandall, personal communication, September 5, 1985, see Appendix A).

For the full form, internal consistency split-half reliabilities were computed separately for the 17 I+ and 17 I- items. Reliability coefficients of .60 were obtained on each scale for older children, and were .54 for I+ and .57 for I- for younger children (Crandall et al., 1965). In the same study, stability coefficients for a sample of 47 children in Grade 3, 4, or 5 yielded reliability coefficients of I_t=.69, I+=.66 and I-=.74, over a two-month period.

New normative data for 1,793 children were reported on the full form and on the shortened form of only 14 items (Crandall, 1978). A test-retest reliability analysis was conducted after a period of 10 days-2 weeks in order to "test the test." Crandall considered the percentage of agreement for each item for the test and retest to be preferable to the phi coefficient because of the uneven split of internal and of external responses. For all 34 items, the average retest agreement after 10 days-2 weeks was 76%; I+=78%, and I-=75%. For the shortened form, the 7 I+ items=83% and the 7 I- items=79% (Crandall, 1978). Point biserial comparisons of the test items to the scale totals for the mean of 17 full scale items was I+=.36 with the full I+ score; and mean of 17 I-=.38 with the full I- score; mean of 7 short form items I+=.48, and mean of 7 I-=.53. These point biserial comparisons are indicators of the internal consistency of the scale (Nunnally, 1970) and indicate greater consistency for the short form. Crandall concluded that both forms "have pretty good psychometric properties. If anything, the short form has a slight advantage" (1978, p. 8.). Others would question whether or not adequate psychometric properties were obtained.

Additional data on reliability from five studies were cited in a handout sent in the IAR packet of materials (obtained from V. C. Crandall, 1985, see Appendix A). These studies, conducted from 1967-1974, cited split-half reliability coefficients (corrected for attenuation) ranging from I+=.47 to .77, I- from .57 to .79, and I_t from .58 to .82. Two additional test-retest reliability studies were cited, at a one-month interval for a sample of 220 eighth graders, I+=.44, I-=.54, and I_t=.58; and at seven-week interval for 248 fourth graders, I+=.48, I-=.60, and I_t=.66.

Split-half and test-retest methods of determining reliability usually give an overestimate of the reliability of the test and, therefore, coefficient alpha is the most generally useful method of determining reliability (Nunnally, 1970). Because of reports of low reliability coefficients, reliability statistics were computed for the I+, I-, and I_t scales for the sample in this study. Split-half Spearman-Brown reliability coefficients, corrected for attenuation, were computed on the IAR pretests for 102 of the gifted children in this sample, I+=.48, I-=.50, and I_t=.52. Coefficient alpha reliability statistics were: I+=-.29, I-=-.13, and I_t=-.50. The 14 items of the short form were analyzed as recommended by Crandall with the split model, Spearman-Brown formula corrected for attenuation: I+=.62, I-=.51, I_t=.49; coefficient alpha: I+=-.12, I-=-.56, I_t=-.72. The coefficient alpha statistics reveal that the IAR scales showed low internal consistency when used in this sample of gifted children.

The test-retest reliability statistics described for the TTCT in Appendix E were computed for the IAR as well as for the other repeated measures in this study. Because of the six-month time interval and intervening significant treatment condition between testing sessions, these reliability coefficients are likely to be an underestimation of reliability of the IAR. The retest reliability coefficient, obtained under these conditions as reported in Table F-9, Appendix F, was I+=.52 and I-=.52.

Instructions for administration and scoring are stated clearly on the copy of the test sent in the IAR packet of materials (obtained from V. C. Crandall, 1985, see Appendix A). It was recommended that individual administration with a standardized tape recorded instruction be used below sixth grade. The test could be administered to a small group of elementary-aged children under taped instructions, but must be carefully monitored. Holen and Newhouse (1976), however, determined that valid results can be obtained from large group (more than 100) administration of the IAR and other attitude scales to junior high school students. Therefore, administration in small groups of 4 to 15 children ought to yield valid results for this study.

The simplified version of the IAR exhibits satisfactory internal consistency and test-retest reliability over a short period of time to justify use in research studies. The short form has the benefit of an empirical demonstration of its construct validity for effort by means of factor analyses. The IAR was chosen for this study, over other possible locus of control measures, despite the reliability coefficients hovering near the borderline range of adequate reliability for research purposes because of its demonstrated good predictive validity. Available research indicated sensitivity to appropriate experimental treatment effects (Bradley & Gaa, 1977; Crandall, 1978; Kanoy et al., 1980; McKenna, 1981) thus indicating validity with appropriate criteria.

> For a test to predict a particular criterion, predictive (criterion related) validity is more important than reliability. When predictive validity is satisfactory, low reliability is not a serious problem (Isaac & Michael, 1981, p. 123).

Moreover, the IAR is specific as an achievement and effort related measure of locus of control.

To provide the benefit of finding the best internal validity, and the stability of the greater length of the test, the full 34-item form was recommended as appropriate in this study with gifted children (V. C. Crandall, personal communication, September 5, 1985). A second analysis was recommended by Crandall (September 5, 1985) to "lift up" the 14 items of the short form to check the empirical validity of the respective scales of the IAR for the present study. The 34-item test was, therefore, used in the present study.

In this study the I+ and I- scores were used in the analyses but the I_t score was not used. Multivariate statistical procedures do not permit use of subscales which contribute to a total scale if an attempt is made to use the subscales and total scale score as part of the same construct. The I_t measure was therefore omitted from the affective construct for the multivariate analyses in this study.

General Self-Concept - Sears Self-Concept Inventory

The measure of general self-concept, Sears Self-Concept Inventory (Sears) (see Appendix M) was to be entered into the multivariate equation first. Based on the review of the literature in Chapter II, it was expected to have a more remote relationship to the creativity measures and to the specific behavioral changes involved in the OM team CPS process than the measure of creative self-concept or of locus of control.

The Sears Self-Concept Inventory: Abbreviated Form is the 1966 revised version. It was made available on the ETS Test Collection of Educational Testing Service, Princeton, New Jersey, November, 1975. In accord with the instructions in the 1966 "Memorandum" appended to the 1975 versions, permission to use the inventory was sought and obtained from Dr. Sears (P. S. Sears, personal communication, July, 1984, see Appendix A).

The Sears inventory is a 48-item five-point Likert-like self-rating scale. (see Appendix M). It taps nine attributes of self-concept: physical ability, attractive appearance, convergent mental ability, social relations with the same sex, social virtues, divergent mental ability, work habits, happy qualities, and school subjects. It is the revised edition of the 100-item scale used in Sears' original validation study (Sears, 1963). Although validation and reliability data are reported for the original scale, there is no manual for the 1966 test and no technical information other than that contained in the "Memorandum" (Sears, 1975).

Kuder-Richardson split-half coefficients of internal consistency reliability were obtained on 32 third graders (Sears, 1975). The coefficients ranged from .56 to .89 for the nine subscales with a total "mental ability" reliability of .92, total "personality" reliability of .88 and total self-concept reliability for all 48 items of .90. The reported internal consistency for the overall tests, combined scales, and those for convergent and divergent mental ability are adequate for use as single scales. The other seven subscales do not, based on the 1966 data, evidence adequate internal consistency to justify use as single predictor scales.

A stability coefficient of the Sears' total score from .12 to .81 with a mean of .50 was reported for 116 third grade students for six intact classes (Sears, 1972) over a four-month interval of time. These data were obtained in the first observational (non-interventional) year of a five-year study of third grade children in three schools of a low-income, predominantly black school district.

The total score for the Sears was selected as the measure to be used in this study. There was insufficient evidence of reliability of the subscale scores to justify use of the subscale scores. The primary reason to utilize a general self-concept scale in the study was to test the hypothesis that a general self-concept scale would be more remote from the effects of OM CPS than a creative self-concept scale, as indicated by the results obtained by Wright et al. (1975) and by analogous results with specific academic self-concept and general self-concept (Shavelson & Bolus, 1982). Therefore, the total score rather than the total of "mental ability" and of "personality" self-concept was selected as the appropriate measure.

The Sears was considered to be an adequate general measure of self-concept for this investigation because it had adequate internal consistency, test-retest reliability, and sensitivity to treatment effects (Whitmore, 1980). The primary reason for selecting the Sears was that the content of the test was appropriate to a study of creativity and a divergent self-concept factor (Sears, personal communication, July, 1985) had been empirically identified. The choice for using the Sears was "strongly supported" by Whitmore who considered the Sears to be "a sensitive instrument that assesses various specific concepts of self" (J. R. Whitmore, personal communication, August, 1985, see Appendix A).

As discussed above, and in Appendix E, test-retest reliability coefficients were calculated for all of the pre- and posttest measures in this study. Because of the six-month time interval and intervening significant treatment effects, as discussed above, these retest reliability correlations are likely to underestimate the stability of the measure. The obtained test-retest reliability is reported in Table F-9, Appendix F, as .69. This was the highest test-retest reliability of any of the measures in the study, and was higher even than the stability coefficient reported by Sears after four months without any treatment conditions. The Sears, then, is a stable instrument that demonstrates good test-retest reliability.

Research Design

This quasi-experimental study utilized three levels of the OM non-repeated factor, and two levels of time, the repeated factor (See Figure 1). Three treatment groups were formed from gifted elementary students (Grades 2-5) of the pull-out enrichment program classes. Group 1 students were those students who did not experience OM treatment. Groups 2 and 3 were students in the program who volunteered to participate on an OM CPS team. Group 2 students' scores on effort on the OM team were below the median for the sample; Group 3 students' scores on effort on the OM team were average or above the median. The independent variable, then, may be described as effort, or degree of participation in the OM team experience.

Limitations

This design was utilized in order to assess the effects of OM CPS team participation on creative and affective scores. The addition of a non-pretest control group was considered unnecessary since a six-month time interval separated the pre- and posttesting. Possible interaction effects of testing and

Groups	N	Pretests	Treatment	Posttests
1. Non-OM	51 49	Creative battery, Affective battery	Enrichment classes; not in OM	Same as pretests
2. OM-lo	32 29	Same	Enrichment classes + OM; OM Effort < median	Same
3. OM-hi	33 33	Same	Enrichment classes + OM; OM Effort $\geq$ median	Same

Figure 1. The Research Design

treatment should be minimized by the time interval and were not considered to be a major threat to the study. A major limitation of causal comparative and quasi-experimental studies is that relationships between cause and effect may be inferred only on a very tentative basis. In the present study, subjects volunteered for the OM condition; therefore, differential motivation between OM and non-OM subjects is a limitation of the study. The use of an intact sample of subjects rather than a random selection is a serious limitation to the generalizability of this study. In order to approach adequate power for the obtained effects for this study, it was necessary to utilize all available subjects for whom permission was obtained. Interpretative analyses of those data which generalize to other populations by inferential statistics should be viewed with some caution because the assumption of random selection of subjects has been violated. The availability of pretest scores prior to the treatment reduces the importance of selection as a limitation to this study.

Moreover, the use of several coaches to train the children may pose the problem of insufficient controls in monitoring the OM training or of ignoring cautions of the limits of coach involvement in the problem solving process (Mansfield et al., 1978; Micklus & Micklus, 1987). The large number of different coaches (n=25), however, would contribute to random effects of coaches and therefore could be considered to contribute to the generalizability of the study.

In addition, the use of multiple pre- and posttest instruments for the study most likely created a reactive effect of experimental procedures which could threaten the external validity of the study. The problem was unexpectedly exacerbated by the administration of the pretest TTCT and affective measures within a few weeks after the annual California Achievement Testing (CAT) sessions. The CAT tests were longer than in previous years. Moreover, the 1985-86 school year was the first year to implement competency testing at 3rd, 7th, and 10th grades in the State of Oklahoma. Third graders had three days of Metropolitan Achievement Tests in the spring, in addition to fall CAT tests. Thus, a contemporary history effect due to the children's feeling they may have been overtested in their regular classes may have inadvertently occurred to threaten the internal validity of the study. The lengthy testing time (approximately 10% of total class time) in relation to the short period of overall class time in the enrichment program caused considerable comment by the children. It is possible that the posttest scores could be lowered due to the childrens' motivation to complete the full test batteries (see Chapter V and Appendix E).

Finally, use of teaching staff to conduct measurements was considered inadvisable by Crandall (personal communication, September 5, 1985). It should also be noted that this experimenter was one of the three faculty members in the enrichment program. However, Isaac and Michael (1981) recommended use of the "regular staff to conduct experimentation within schools, especially where findings are to be generalized to other classroom situations, to make classroom research nonreactive, i.e., to reduce the Hawthorne or guinea pig effect." (p. 91) Hawthorne effects were controlled in that the sample of all students for whom data was collected, whether OM or control, all knew that they were in a study, but had no way to know whether they were in an experimental or control group.

Procedure

Curriculum in the Enrichment Program

Enrichment program students meet in pull-out classes from their regular classes one-half day per week in each of the ten elementary schools (see Appendix C). The OM program had been available to students in the enrichment program as an extra-curricular option for the past six years.

Eligible gifted students are admitted at various times during the year as they qualify on the basis of individual testing. The program pull-out classes vary in size from 4-15+ students. Most classes are cross-grade grouped in eight of the ten schools, with Grades 1-5 in the same class in smaller schools and only 1-3 and 4-5 together in other schools. In general, the program primarily serves grades 2-5 with only a few students entering the pull-out program during first grade.

The children are taught by teachers who are specialists in gifted education. Two of the three itinerant teachers have each taught gifted students for ten years and have master's degrees in special education related areas and numerous workshop and graduate experiences in gifted education. The third teacher, new to the program in 1985-1986, was completing her master's degree in gifted education.

All three treatment groups experienced the basic enrichment program in an accepting environment as part of their two to two and one-half hour per week class for the mentally gifted. The goals of the program are to develop critical and creative thinking skills, feel positive about oneself and others, communicate effectively, employ research skills, and explore, study, and share areas of interest. Included among the many type II process training activities (Renzulli, 1977) is use of materials from "New Directions in Creativity" (Renzulli, 1973). Students experience several areas of choice in meeting their individualized contracts with the teachers. All students work critical thinking problems, such as in logic, and creativity problems and do some form of independent research project. The varied activities in this process-based curriculum are described in Appendix D.

All students experience training and practice in the CPS team process as part of the Future Problem Solving program during class time. The steps in the Future Problem Solving CPS process are: 1) researching the topic, 2) brainstorming problems, 3) identifying the underlying problems, 4) brainstorming solutions, 5) selecting criteria, 6) evaluating solutions, and 7) describing the best solution (Crabbe, 1985).

Fourth and fifth grade students participate in all seven steps of the Junior Future Problem Solving Program, but only work two of the three problems included in the material from the national program. They are trained and practice on steps one through three for practice problem number one and steps one through seven of practice problem two (Crabbe, 1985). Second and third grade children have briefer experiences as part of the Primary Future Problem Solving Program (Crabbe, 1985). (See Appendix N for a copy of sample problems.) They are trained and practiced on step three for the first practice problem and steps one through three for their second problem. All students experience a training problem to teach team evaluation skills of generating criteria and consensual ranking of choices. Any student entering after March of any year would have missed the major CPS work for that year, and would have missed the effects of such training.

Procedure for the Study

All students for whom parental permission was obtained were used in this quasi-experimental research study. The available population consisted of a total 116 for the creative sample (111 for the affective sample) who were distributed in the three naturally occurring treatment groups as follows: non-OM=51(49), OM-lo=32(29), and OM-hi=33(33).

Prior experience in OM was measured by the number of years the child participated on an OM team. Experience in the enrichment program was measured by the number of months that the child was enrolled in that enrichment program. In that enrichment students volunteer to be on an OM team, participation on a team can vary independently of their enrollment in the program. Students also vary, within their grade level, with the length of time in the program.

The first measures to be administered for the study were the figural portion of the TTCT and Similes which were group administered in mid-October to students during their enrichment class time. The verbal battery of the TTCT was administered, in general, the following week as a group to all fourth and fifth grade students during their enrichment class sessions. The TTCT verbal battery was individually administered over a four-week period to second and third grade students in the study in late October-mid November. Individual administration was performed by graduate students in counseling or testing or by the experimenter during enrichment class time or after school.

Care was taken to provide consistent warm-up activities (Appendix O) and consistency of the general motivating atmosphere among the various classes for administration of creative and affective and pre- and posttests. (See Appendix O for sample warm-up activities and sample instructions for administration of the various measures.) A five-minute figural class exercise was used as the warm-up for the figural and Similes pre- and posttests. A verbal brainstorming task was used for approximately three minutes with two minutes of sharing of ideas to precede the verbal TTCT tasks. On the warm-ups, the administrators asked the class or individual to look for additional ideas, and encouraged them to think of "what else, or what could you add?" An affective warm-up (see Appendix O) was administered just preceding the first of the affective tasks, Ideas. Students were asked to evaluate their typical style of going about daily work and were welcome to briefly share their responses if desired.

The pretests in self-concept, creative self-concept and locus of control were administered in November to those students who were permitted by their parents to be tested for this study. In some classes, pretesting of the last measure, the IAR, was not completed until mid-December. Ideas was read aloud to second and third graders by the teacher. The IAR was taken by students listening to taped instructions and reading of each item.

The first of two Future Problem Solving problems, in general, spanned about a three-week period for Juniors, and one to two weeks for Primary level teams in November. Although the national program provides three problems at each level, because so much effort in the district's enrichment program is involved in OM team work by many of the children and the teachers, the program only provides two of the Future Problem Solving problems. The second problem involved about a three to four week period for research and problem solving for Juniors and one to two weeks for Primary students, ending in mid-January. The teachers evaluated the children's performance on their Future Problem Solving teams on the Teacher Rating Scale of Team Effectiveness (see Appendix I), usually within two weeks after the team completed their work.

In November all students in the program were shown slides of the previous year's local OM competition. During their enrichment classes, portions of the OM problems were read and discussed in class in mid-October and students were encouraged to sign up for a team. National guidelines for OM suggest that teams can be formed by having students try out in order to select the most creative. In the school district under study, the participation in OM is open to elementary and secondary students who are identified as gifted or enrolled in a secondary honors program. To a large extent students choose their own team members. Teams are formed based upon interest in being in OM, interest in first or second choice of a specific OM problem, and to some extent upon consideration of friendship or classmate groups. The coordinating enrichment teachers facilitate and approve formation of teams. In order to form the greatest number of teams and to maximize inclusion of all enrichment children who expressed an interest in being on a team, the teachers helped five teams to form where the children attended different schools in the district.

Teams met at home after school under the guidance of a trained volunteer coach from December-January until the competition on March 1, 1986. Most teams met at least six times for about two-hour sessions. Although national guidelines (Micklus & Micklus, 1987) suggest that teams should consist of five to seven members, for local competition three teams consisted of only three or four students. The coach, a volunteer parent, attended a two-hour training session conducted by the enrichment program teachers to give some familiarity with the rules of OM and facilitation of the CPS process, including techniques of brainstorming (see Appendix D for further description of the CPS process in OM). All students, OM and non-OM, received intensive practice in brainstorming techniques in enrichment classes during the three weeks immediately prior to the competition. The local competition was held on March 1 with 83 elementary children on 17 teams in four different problem areas. Of these 83, 65 were children under study. Of the remaining 18, 14 were students who started in the program some time between early October and mid-January, and therefore could not be included in the study. Two students moved before complete posttest data could be obtained, and two were children for whom parents did not give permission to be tested (see Table F-5, Appendix F). Another 21 middle-school students from four teams competed at the local meet.

The Rating Scale of Work Done by the Team was administered during the week immediately following the local competition. The Student Evaluation of OM was administered two weeks prior to the March 1 competition and again during the week after the local OM meet. Questions were read aloud for administration. The Student Evaluation of OM was administered the second time to assess possible effects of the experience of competition and knowledge of the judges' ratings on self-perception of work done by the team for possible future analysis of the effects of competition.

Subjects were assigned to groups OM-lo or OM-hi based on their score on the Rating Scale of Work Done by the Team. Those who scored below the median, 10.3, were considered to be in the low effort in OM group (OM-lo). Those subjects whose participation and contributions to the team were scored at the median or higher were assigned to the OM-hi group. The posttests in creativity and affect, and Student Evaluation of OM, were administered in late April or early May. Second and third grade children were posttested individually on the verbal TTCT during that period with the exception of make-up tests. All testing was completed prior to the last three weeks of school.

Procedure for Missing Data

The Rating Scale of Work Done by the Team (Appendix H) was completed by all students who were in OM in class the week following the competition. However, during data analysis it was noted that nine children, distributed on 4 of 17 teams, had not completed the rating scale of student perceived effort necessary to assign the subjects to the independent variable grouping of high or low effort in OM. All of these children attended the enrichment program, but were not in the study as explained above. Two were students whose parents had not given permission for them to participate in the study. One moved out of the district the week following OM. The other five were students new to the program for whom permission to complete these OM forms had not been obtained. The ninth child was in the study but was often uncooperative about taking the tests.

In early July, the parents of all nine of the children were contacted. The parents were highly cooperative and in most cases let the experimenter test the child at their home. The child who was most uncooperative during class time had been rebelling because his mother had given permission for him to be in the study without consulting him. When the researcher asked the child if he would be willing to fill the form out and explained why it was important, he was then quite willing to do so. Seven of the missing effort scales were thus completed. The remaining two forms were from two children who were on OM teams but did not participate in the study. Their effort forms were not completed so scores were estimated from the rankings of their teammates in order to calculate the median effort score for the entire sample. Their estimated effort scores were excluded from calculating effort scores for the remaining members of their teams.

There were an additional seven students for whom the affective posttest battery data was incomplete. These parents and children were also contacted during July, and the missing tests were completed at that time. Pretest data that was not completed by December was left incomplete.

It was noted during data analysis that three subjects had only partially completed the last page of the Sears test. Because 75% of the test was already completed, the total score was estimated for these children based upon responses to the portion of the test already completed. This procedure is recommended by

Tabachnick and Fidell (1983) to prevent unnecessary loss of data. Total score estimation was used for the four children who left 1-2 items blank on Ideas. On the IAR, however, because different subscales were to be completed, no attempt was made to estimate scores. Seven students omitted 2 or 3 items on the IAR, and one unintentionally did not respond to five items. Scores for these eight were considered to be complete without any estimation for the missing items. If any of the four affective measures was less than 75% complete, the test was considered to be too incomplete to be counted.

Procedure for Analysis of the Data

The data were analyzed using two separate multivariate analyses. The first was a 3 x 2 multivariate mixed model analysis of variance (MANOVA) with the independent variables of effort in OM (OM-lo, OM-hi, non-OM) and time (pre, post). The 10 dependent variables forming the construct of creativity were the responses on Similes; TTCT verbal: standard scores for the three norm-referenced measures, fluency, flexibility, originality; TTCT figural: standard scores for the five norm-referenced measures, fluency, originality, abstractness of titles, elaboration, resistance to premature closure, and a raw score for the total of the creative strengths.

The second analysis examined the four dependent variables forming the affective construct: creative self-concept, 1) Ideas; locus of control, 2) I+ and 3) I-; and 4) general self-concept, Sears. The raw scores were used as the dependent score values for these measures.

A preliminary inspection of the data was made to determine if they met the assumptions of analyses of variance and multivariate analyses of variance (see Appendix B, Procedural Issues of Multivariate Statistics). Outliers among covariates were checked for use by SPSS-X Regression (SPSS, 1986) and found not to contribute to heterogeneity of regression. Some problems with possible violations of these assumptions were encountered, but judged to be acceptable (see discussion in Appendix B).

A priori, the following analytical procedure was determined: If either analysis showed a significant global main effect, the following post hoc strategy would be used to interpret the data. A strength of association of the global analysis and of the large univariate dependent variables would be performed. If the three-level factor, effort in OM, yielded a significant global main effect, the post hoc tests for specific comparisons would be calculated. Simple Scheffe contrasts would be performed on those univariate variables that were significant to determine which levels of the OM factor were accounting for the differences in the scores. In the affective MANOVA, the stepdown F tests would be examined to determine if the variables of greatest interest, creative self-concept, and then locus of control accounted for a significant portion of variance between groups when examined in their purest manner. In the creative MANOVA, the stepdown F tests would be examined to determine if the verbal measures of fluency, flexibility, and originality accounted for a significant portion of variance between OM groups when examined in their purest manner, as indicated by prior studies (Rose & Lin, 1984).

To assess the effect of prior experience in OM and differing amounts of previous experience in the enrichment program, two sets of supplemental analyses were to be performed. An analysis of the effect of previous experience in the program on both sets of dependent variables was to be computed: A 3 x 2 x 2 MANOVA with the independent variables of effort in OM (non-OM, OM-lo, OM-hi), experience in the enrichment program (less than 13 months of experience, 13 or more months of experience), and time (pre, post). A 3 x 2 x 2 mixed model MANOVA was also to be performed with the independent variables of effort in OM (non-OM, OM-lo, OM-hi), prior experience in OM (no prior experience, prior experience on an OM team), and time (pre, post). These variables were considered in analyzing both sets of data, affective and creative.

The hypothesis-wise alpha level was set at .05 for all comparisons in the study.

Plan of Analyses to Test the Hypotheses

The first ten hypotheses offered in Chapter I each predicted that a significant interaction would occur.

1. There will be a significant interaction on the creativity construct scores with respect to effort in OM and time.

2. There will be a significant interaction on the affective construct scores with respect to effort in OM and time.

3. There will be a significant interaction on the creativity construct scores of students in relation to the amount of their experience in the enrichment program and time.

4. There will be a significant interaction on the affective construct scores of students in relation to the amount of their experience in the enrichment program and time.

5. There will be a significant interaction on the creativity construct scores of students in relation to the amount of their prior experience in OM and time.

6. There will be a significant interaction on the affective construct scores of students in relation to the amount of their prior experience in OM and time.

7. There will be a significant interaction on the creativity construct scores when these scores are compared with respect to effort in OM, experience in the enrichment program, and change in scores from pretest to posttest in the experimental year.

8. There will be a significant interaction on the affective construct scores when these scores are compared with respect to effort in OM, experience in the enrichment program, and change in score from pretest to posttest in the experimental year.

9. There will be a significant interaction on the creativity construct scores when these scores are compared with respect to effort in OM, prior experience in OM, and change in score from pretest to posttest in the experimental year.

10. There will be a significant interaction on the affective construct scores when these scores are compared with respect to effort in OM, prior experience in OM, and change in score from pretest to posttest in the experimental year.

11. There will be a significant difference in sensitivity of the scores of the univariate dependent variables of the affective construct to student effort in OM.

12. There will be a significant difference in creative self-concept scores with respect to student effort in OM and time.

Six mixed model MANOVA analyses were planned to test the twelve hypotheses. These analyses and their relation to the hypotheses are:

1. A 3x2 mixed model MANOVA with the between groups variable of OM effort (non-OM, OM-lo, OM-hi) and the within group variable of time (pre, post) was used to analyze the creative construct for hypothesis 1. The interaction between OM and time was the effect for the test of hypothesis 1.

2. A 3x2 mixed model MANOVA with the between groups variable of OM effort (non-OM, OM-lo, OM-hi) and the within group variable of time (pre, post) was used to analyze the affective construct for hypothesis 2. The interaction between OM and time was the effect for the test of hypothesis 2.

Inspection of the outcome of the univariate and stepdown analyses of the four dependent variables comprising the affective construct for the interaction of OM and time was the test of hypothesis 11. The interaction effect between OM and time for the univariate dependent variable of creative self-concept was the test of hypothesis 12.

3. A 3x2x2 mixed model MANOVA with the between groups variables of OM effort (non-OM, OM-lo, OM-hi) and experience in the enrichment program (less than 13 months experience, 13 or more months experience) and the repeated variable of time (pre, post) was used to analyze the creative construct to test hypotheses 3 and 7. The interaction between experience in the enrichment program and time is the test of hypothesis 3, and the three-way interaction between experience in the enrichment program, OM effort, and time is the test of hypothesis 7.

4. A 3x2x2 mixed model MANOVA with the between groups variables of OM effort (non-OM, OM-lo, OM-hi) and experience in the enrichment program (less than 13 months experience, 13 or more months experience) and the repeated variable of time (pre, post) was used to analyze the affective construct to test hypotheses 4 and 8. The interaction between experience in the enrichment program and time is the test of hypothesis 4, and the three-way interaction between experience in the enrichment program, OM effort, and time is the test of hypothesis 8.

5. A 3x2x2 mixed model MANOVA with the between groups variables of OM effort (non-OM, OM-lo, OM-hi) and prior experience in OM (no prior experience on an OM team, prior experience on an OM team) and the repeated variable of time (pre, post) was used to analyze the creative construct to test hypotheses 5 and 9. The interaction between prior experience in OM and time is

the test of hypothesis 5, and the three-way interaction between prior experience in OM, OM effort, and time is the test of hypothesis 9.

6. A 3x2x2 mixed model MANOVA with the between groups variables of OM effort (non-OM, OM-lo, OM-hi) and prior experience in OM (no prior experience on an OM team, prior experience on an OM team) and the repeated variable of time (pre, post) was used to analyze the affective construct to test hypotheses 6 and 10. The interaction between prior experience in OM and time is the test of hypothesis 6, and the three-way interaction between prior experience in OM, OM effort, and time is the test of hypothesis 10.

Summary

The methods for the study were presented in this chapter. The characteristics of the subjects and the seven instruments were discussed and the research design and procedures were explained.

The present study, therefore, addresses a needed area of research. It is quasi-experimental in design so that tentative influences about causality may be discussed. It utilizes multivariate statistics which can illuminate the variables and determine the degree of purity of their relationships to the independent variable(s) under investigation. Most important, the present study utilizes these appropriate research methodologies in an area of research that to date has insufficient study, that of effects on creativity of team CPS in gifted children (Crabbe, 1985; Hendrickson, 1985/1986; Micklus & Micklus, 1987; Torrance, 1972a, 1985). It is hoped that the present study will establish a measure of criterion validity for the streamlined scoring of the TTCT in relation to a new and needed criterion, the skills involved in the CPS process.

CHAPTER IV

RESULTS

Overview

The data were analyzed by two sets of three mixed model multivariate analyses of variance. The first set tested the creative constructs for the primary hypothesis for the effect over time of participation on an OM team. Prior experience on an OM team and previous experience in the enrichment program were tested for their effect on change in scores over time on the dependent variables in relation to student participation in OM in two supplemental analyses. The second set tested the affective constructs for the primary hypothesis of OM x Time and then the hypotheses concerning experience in the two supplemental MANOVAs.

All analyses were performed with SPSS^{X} MANOVA (SPSS, 1986) with the hierarchical (default) adjustment for nonorthogonality. The data were tested to determine satisfaction of the assumptions necessary to perform repeated measures MANOVA and these results are addressed as procedural issues in Appendix B. Analysis of the data revealed a need for alternative methods of analysis in addition to the aforementioned two sets of three analyses. Differences between grade levels were found to affect variables of the creative construct, therefore, a MANOVA testing the Grade x OM was performed on the creative construct. Assessment of the preliminary analyses reported in this section indicated a need for two transformations of the creativity scores and for use of covariance analyses for both constructs. The rationale for the transformation of setting the ceiling of 160 for extremely high TTCT scores was to reduce the impact of outlying scores and of regression toward the mean of such scores on posttest analyses (see p. 145). The rationale for the transformation of the ten dependent variables to an abbreviated creative construct consisting of three dependent variables was to reduce the effect of excessively high correlations among verbal creativity variables. These procedures are explained below (see p. 157) under possible violation of assumptions of multicollinearity and in Appendix B (see test of the assumptions for MANOVA and MANCOVA, Appendix B). To control for the pre-existing individual differences among subjects within the cells, multivariate analyses of covariance (MANCOVA) were also performed on the creativity data and on the affective data to test for differences between OM effort groups on the adjusted posttest scores.

Tests of Effects on the Creative Construct

Multivariate Analyses of the Creativity Data

The ten dependent variables forming the construct of creativity were measured by three instruments: Similes, which provided one raw score; TTCT verbal, which provided three standard scores (fluency, flexibility, and originality); and TTCT figural, which provided five standard scores--fluency, originality, abstractness of titles (titles), elaboration, and resistance to premature closure (closure)--and a raw score for total indicators of creative strengths. The ordering of the dependent variables was set in an a priori hierarchy to examine the scores on the verbal TTCT in a pure manner, that is, with removal of variance contributed by the other dependent variables. Therefore, the figural test scores were entered first into the equation, then Similes, and lastly the verbal TTCT scores in the following order: figural fluency, figural originality, titles, elaboration, closure, total of creative strengths, Similes, verbal fluency, flexibility, and last, verbal originality.

Examination of the within cells correlation matrices of the dependent variables for creativity showed that a multivariate construct was formed for all analyses. Each of the ten dependent variables shared a Pearson correlation of greater than .30 with at least one other variable (see Table P-1 in Appendix P), where the correlations for pretest variables ranged from -.19 to .95. The verbal TTCT measures were very highly correlated with each other, .91 to .95. The total of creative strengths variable showed correlations greater than .30 with figural titles, elaboration, and closure, and with Similes. Values of correlations were lower on posttest scores, ranging from -.22 to .84.

Effect of Odyssey of the Mind Effort:

Analysis Prior to Adjustment of Scores

A 3x2 mixed model MANOVA was performed on the ten creative dependent variables. The fixed independent variables were OM (non-OM, OM-lo, OM-hi) for the between groups comparison and time (pre, post) for the repeated measure. Cell means and standard deviations are reported in Table P-2 (Appendix P). For purposes of providing a baseline for comparison of the effect of transforming the scores a 3x2 mixed model MANOVA was performed on the original scores. As seen by the analysis reported in Table P-3 (Appendix P) the results are comparable to the 3x2 mixed model MANOVA where the TTCT standard scores were transformed by setting a ceiling, a maximum possible score of 160 (Table 1). The differences between these analyses will be discussed in Chapter V.

Adjustment of Ceiling for TTCT Scores

Examination of the cell means reveals that the sample of gifted children in the study displayed remarkable evidence of high cognitive creativity. The mean scores for the verbal activities and for figural elaboration are one and one-half

Table 1

Summary of Mixed Model Creative MANOVA for Odyssey of the Mind (OM) Effort: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Ten Creative Dependent Variables[a]

Source	Multivariate Global F	Stepdown Fs/Univariate Fs — Dependent Variables: Similes	Verbal TTCT: Fluency	Verbal TTCT: Flexibility	Verbal TTCT: Originality
Between Subject Analyses					
OM	1.84*[b]	3.50*[c] 8.84***	1.67 .75	.36 .36	1.12 .73
Repeated Subjects Analyses					
Time	30.96***	28.21*** 28.92***	7.63** 7.76**	.72 .14	43.88*** 88.40***
OM x Time	1.27	2.66 2.34	1.42 2.74	1.94 3.75*	.40 3.34

Source	Figural TTCT: Fluency	Originality	Titles	Elaboration	Closure	Creative Strengths
Between Subjects Analyses						
OM	1.19 1.19	1.74 2.77	.45 .35	1.60 2.32	4.62** 7.54***	2.24 7.15***
Repeated Subjects Analyses						
Time	13.44*** 13.44***	63.58*** 84.31***	2.60 2.45	11.56*** 26.28***	.48 2.02	3.14 3.03
OM x Time	1.54 1.54	1.13 1.13	.20 .27	.55 .37	1.47 1.55	1.34 .62

[a]All standard scores set at ≤ 160.
[b]Pillai's criterion for statistical inference.
[c]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001

standard deviations above the norm of 100 thus indicating that the sample of children in this study is highly creative. (The TTCT provides standard scores with a mean of 100 and standard deviation of 20, see Appendix E.) In one cell, non-OM, the mean scores for verbal flexibility were a remarkable 136.24, close to two standard deviations above the mean for their age group (see Table 2). (Prior to adjusting scores at the ceiling of greater than or equal to 160, the mean was 141.78, see Table P-2, Appendix P). It should be noted, moreover, that the mean standard scores for the pretest flexibility variables (137.27) are greater than seven points higher than verbal fluency (128.97) or verbal originality (130.19). The fluency and originality scores, although extremely high, actually do not reflect the level of creativity in this sample. As stated in Chapter III, p. 110, only six of the seven TTCT verbal activities were administered. The activity which was omitted, Activity 6, Unusual Questions, unlike the other six activities, is scored only for fluency and originality (Torrance, 1974a). Therefore, flexibility, based on the six activities contributing to verbal flexibility, is the truer comparison of this sample's verbal creativity in comparison to the normal population. The mean verbal creativity of this sample, based on six subtests, 132.14 prior to adjustment of the scores, as measured by the TTCT, then, is over 1½ standard deviations above the normal population.

Examination of the verbal TTCT scores showed that a few students in the sample were profoundly gifted in verbal creative thinking. Seventeen students earned standard scores on the mean of the three measures of the verbal TTCT at greater than or equal to 160, more than three standard deviations above the norm. Of these, five had scores in excess of 200. In all of these cases, there was very significant regression toward the mean upon posttesting, with the most pronounced in a third grade student who was not experienced in OM, but experienced in the enrichment program, verbal originality pre=231, post=116. To

Table 2

Cell Means and Standard Deviations of Ten Creative Dependent Variables and Odyssey of the Mind (OM) Effort[a]

		Similes		Verbal TTCT Dependent Variables					
				Fluency		Flexibility		Originality	
OM Effort	n	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Pretest Means									
Non-OM	51	37.92	13.64	127.67	22.78	136.24	19.16	129.12	21.83
OM-lo	32	46.53	11.27	124.19	19.48	132.97	15.72	125.16	19.00
OM-hi	33	50.24	21.21	124.46	20.22	130.91	15.88	125.27	18.34
Posttest Means									
Non-OM	51	47.67	16.16	119.82	17.85	132.47	18.56	108.86	14.74
OM-lo	32	49.44	14.41	116.13	19.45	130.00	18.32	106.63	18.40
OM-hi	33	61.12	18.21	126.18	20.75	137.46	20.28	115.76	17.36

		Figural TTCT Dependent Variables											
		Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
		$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Pretest Means													
Non-OM	51	98.39	16.22	114.73	20.03	110.77	17.39	124.73	19.70	105.88	13.01	8.49	2.94
OM-lo	32	104.31	14.30	124.31	17.09	109.00	13.67	132.97	19.18	111.97	9.33	9.47	2.21
OM-hi	33	105.64	18.94	124.46	19.88	111.91	21.05	133.39	19.00	110.46	11.16	10.15	3.25
Posttest Means													
Non-OM	51	107.69	15.21	135.14	20.89	112.24	17.69	116.29	19.85	105.82	17.01	8.84	2.82
OM-lo	32	110.03	14.91	137.69	22.24	113.63	17.95	120.88	22.39	113.59	13.82	9.75	2.98
OM-hi	33	107.76	18.58	142.15	17.61	115.97	20.22	121.64	19.69	117.70	17.26	11.24	3.35

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at $\leq$ 160.

control for the inordinate effect of these extremely high scores on the overall analyses, all subsequent analyses were therefore performed on scores transformed by setting the ceiling of any standard TTCT score at 160. It was recommended by O. Anderhalter that there was no need to use scores more than three standard deviations above the mean of the norm group (personal communication, June 9, 1988, see Appendix A). Although the mean score of the sample was closer to 130 than 100 this transformation of scores was validated by reducing the variability without such transformation. As may be seen by comparing the standard deviations as reported in Table 2 and in Table P-2 in Appendix P, there was greater variability among the cell means prior to transformation of the scores. The scores were transformed in order to reduce the impact of outliers by moving extreme cases more toward the measure of central tendency as recommended by Tabachnick and Fidell (1983).

Remarkably high verbal creativity is demonstrated as well by the mean pretest score on Similes of 43.80 for this sample of 116 Grade 2-5 gifted students. Means and standard deviations for nine different fourth and fifth grade groups were presented in the manual (Schaefer, 1971). These means ranged from $\overline{X}$=20.70, SD=12.30 for 105 fourth grade children from a school serving a disadvantaged area to $\overline{X}$=47.83, SD=12.59 for 65 fifth grade children from a school serving an upper middle class area. Percentile scores were provided for a group of 130 middle class fifth grade students. The mean of 43.80 for the 116 students in grade 2 through 5 places the children in this study at the 82nd percentile of children who were older than the mean age of those in this sample. The mean for the fifth grade students in this sample was 52.79, higher than the mean of any of the norm groups, and at the 95th percentile for the select normative group.

Effect of Odyssey of the Mind Effort:

Adjusted Scores

A 3x2 mixed model MANOVA was performed on the ten creative dependent variables where the TTCT maximum possible score was set at a standard score of 160. The purpose of this analysis was to test hypothesis 1, the major hypothesis of the study, the OM x Time interaction. The independent variables were effort in OM (non-OM, OM-lo, OM-hi) to compare effects between groups, and time (pretest, posttest) to compare changes within individuals. The cell means and standard deviations are reported in Table 2.

The results of the MANOVA for OM effort showed a significant difference on the multivariate creative construct in pretest scores for OM groups judged by Pillais' criterion, $F(20,210)=1.84$, $p < .05$, and between time of testing, $F(10,104)=30.96$, $p < .001$ as shown in Table 1. The results reflected a strong association between the effects of time (pre to posttest scores) and the combined dependent variables, $N^2=.75$. The relation of OM to the construct was less marked, $N^2=.28$. (For a description of N^2, the estimate of strength of association, see Appendix B.) The interaction testing the major hypothesis, OM x Time, did not yield a significant difference for the global multivariate construct. However, because main effects that were not predicted for OM and for time were found to yield significant global results, it is appropriate to consider the significant univariate F statistics associated with the ten dependent variables for these effects. These univariate statistics are reported in Table 1.

The pre-existing differences between the OM effort groups are shown by the significant univariate statistics (all df=2,113; all $p < .001$): Similes, $F=8.84$, $N^2=.11$; closure, $F=7.54$, $N^2=.07$; and total of creative strengths, $F=7.15$, $N^2=.08$ The analyses also resulted in a significant stepdown F for two of the three measures: creative strengths, $F(2,109)=4.62$, $p < .01$; and Similes, $F(2,107)=3.50$, $p < .001$. The decreasing values of the degrees of freedom for the total number of observations is noted in stepdown analyses because the addition of another dependent variable at each step accounts for a degree of freedom which must be removed. The pretest means for the non-OM group are lower on these measures than those for students who chose to be on an OM team. On these three variables, scores were lowest for the non-OM group. Pretest scores on total of creative strengths and Similes were higher for the OM-hi group than for OM-lo or non-OM. These results are shown graphically in Figure 2. Post hoc comparisons using simple contrasts and $p < .05$ as the level of significance revealed a strong difference in verbal creativity as measured by the Similes test between those students who were on an OM team and those who were not, such that OM-hi students had significantly higher scores than non-OM ($t=-4.20$, $p < .001$) and OM-hi were also higher than OM-lo ($t=-2.26$, $p < .05$). On ability to resist premature closure, OM-hi students were significantly stronger than non-OM ($t=-3.50$, $p < .001$). Students volunteering to work on an OM team were stronger in the total for creative strengths than those not in OM, with OM-hi significantly higher than non-OM ($t=-3.77$, $p < .001$). These results indicate that the groups were not equivalent in scores on creativity measures prior to experiencing the experimental treatment, effort on an OM team, during the year of the study.

There were significant changes over time in scores on measures of the overall creativity construct $F(10,104)=30.96$, $p < .001$, $N^2=.75$. Support for the construct was found in the univariate analyses (all $df=1,113$; all $p < .001$): verbal originality, $F=86.55$, $N^2=.21$; figural originality, $F=84.28$, $N^2=.20$; Similes, $F=28.37$, $N^2=.06$; figural elaboration, $F=25.86$, $N^2=.07$; figural fluency, $F=13.17$, $N^2=.04$; and verbal fluency, $F=7.62$, $p < .05$, $N^2=.01$. The stepdown analyses resulted in a significant F for all six of the variables: figural originality,

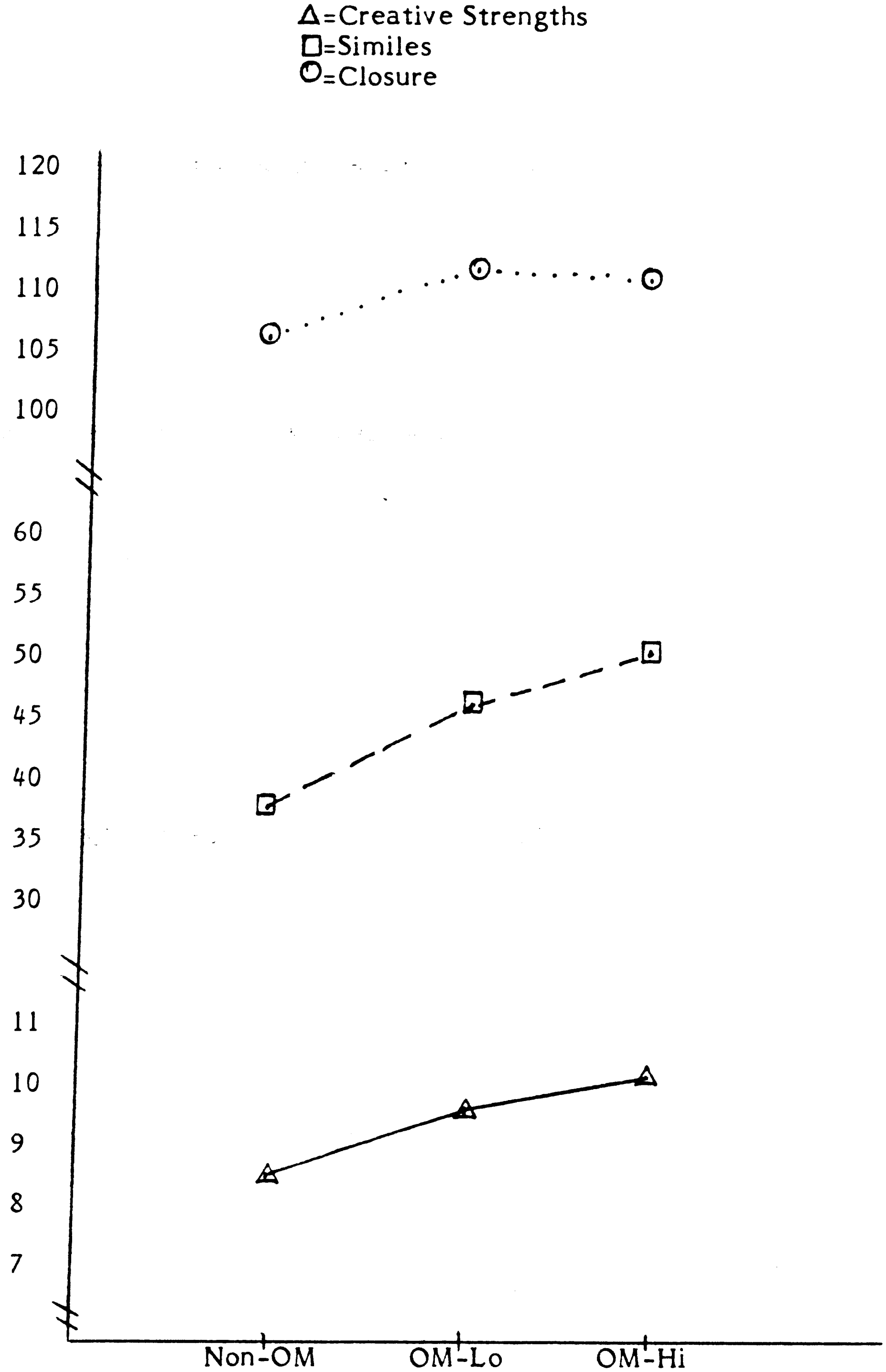


Figure 2. Pre-existing Differences Between Odyssey of the Mind (OM) Effort Groups of Pretest Scores

Note. Creative Strengths and Similes are reported in raw scores; Closure is reported in standard scores set at ≤ 160.

$F(1,112)=63.58$; verbal originality, $F(1,104)=43.88$; Similes, $F(1,107)=28.21$; figural fluency, $F(1,113)=13.44$; and figural elaboration, $F(1,110)=11.56$, all at $p < .001$, and verbal fluency, $F(1,106)=7.63$, $p < .01$. Inspection of the means shows an overall increase in scores from pre- to posttest for figural fluency, and a very marked increase for Similes and figural originality. Scores, however, decreased on figural elaboration and verbal fluency and very markedly on verbal originality. The differences in the pre- and posttest scores in which there were significant findings may be more readily understood by examining the means. The means of the following variables increased over time: figural originality (pre=120.14, post=137.84); figural fluency (pre=102.09, post=108.35); Similes (pre=43.80, post=51.98). The variables which decreased over time were: verbal originality (pre=126.93, post=110.21); figural elaboration (pre=129.47, post=119.08); and verbal fluency (pre=125.79, post=120.61).

Significant main effects which had not been predicted were found for differences between OM effort groups in their pretest scores and between pretest and posttest scores across groups. These results indicate that the OM effort groups were not equivalent in pretested creativity, such that children who chose to join an OM team in January showed greater tested creativity in November than children who did not choose to join OM. The decrease in verbal creativity scores of the TTCT was contrary to the results expected from the review of the literature. Normally, an increase in posttest scores reflects creativity training. As discussed in Chapter V and Appendix E, a context effect appears to have significantly contributed to depression of the posttest scores.

The primary hypothesis that there would be a difference between creativity test scores among the different OM effort groups in change over time was not supported.

Effect of Experience in the Enrichment Program

A 3x2x2 mixed model MANOVA was performed on the creative construct to test hypothesis 3, the effect of the interaction of the amount of experience in the enrichment program and time, and hypothesis 7, the interaction of OM x Experience in Enrichment Program x Time. The independent variables were OM (non-OM, OM-lo, OM-hi) and experience in the enrichment program (less than 13 months experience, 13 or more months experience) to compare effects between groups, and time (pre, post) to test for changes within individuals. Observed cell means and standard deviations are presented in Table P-4, Appendix P.

A summary of this 3x2x2 MANOVA is reported in Table P-5, Appendix P. As already reported in the first creative analysis, a significant main effect was found for pre-existing differences between OM groups and for change in scores over time across groups. The analysis failed to support either of the hypotheses for the interactions, hypothesis 3 and 7, thus indicating that differences in amount of prior experience in the enrichment program did not differentially affect change in creativity over time.

A significant multivariate pretest difference was found for experience in the enrichment program, $F(10,101)=3.13$, $p < .05$, $N^2=.24$. Support for the construct was found in the univariate analyses (all df=1,110): total of creative strengths, $F=12.06$, $p < .001$, $N^2=.06$; figural closure, $F=7.88$, $p < .01$, $N^2=.04$; figural elaboration, $F=6.45$, $p < .05$, $N^2=.04$; Similes, $F=6.50$, $p < .01$, $N^2=.03$; verbal flexibility, $F=4.35$, $p < .05$, $N^2=.04$; figural originality, $F=5.87$, $p < .05$, $N^2=.03$. The analyses also resulted in a significant stepdown F on four of the preceding six measures (all at $p < .05$): figural elaboration, $F(1,107)=5.76$; verbal flexibility, $F(1,102)=4.49$; closure, $F(1,106)=4.45$; and total of creative strengths, $F(1,105)=4.11$. A comparison of the means for these respective measures reveals that with the exception of verbal flexibility, less than 13 months=137.52, greater than or equal to 13 months=131.87, the students with 13 or more months of experience in the enrichment program showed greater evidence of cognitive creativity than students with less experience in the program. These means were: total of creative strengths, less than 13 months=8.45, greater than or equal to 13 months=9.64; closure, less than 13 months=106.05, greater than or equal to 13 months=110.34; figural elaboration, less than 13 months=123.47, greater than or equal to 13 months=132.62; Similes, less than 13 months=38.65, greater than or equal to 13 months=46.52; and figural originality, less than 13 months=116.40, greater than or equal to 13 months=122.11. These results, therefore, might lend strength to previous findings cited in the literature that creativity can be enhanced by creativity training.

Prior Experience in Odyssey of the Mind

A 3x2x2 mixed model MANOVA was performed on the creative construct to test hypothesis 5, the effect of the interaction of prior experience on an OM team and time, and hypothesis 9, the interaction of OM effort x Prior OM x Time. The independent variables were OM (non-OM, OM-lo, OM-hi) and prior experience on an OM team (no prior experience, prior experience on an OM team), and time (pre, post). Cell means and standard deviations are presented in Table P-6, Appendix P.

A summary of this 3x2x2 MANOVA is reported in Table P-7 of Appendix P. As previously reported in the first analysis of OM effort and the creative construct, significant differences were found across groups over time. The analyses did not support either hypothesis 3 or 7, the hypotheses for the interactions, thus indicating that differences in amount of previous experience on an OM team did not affect the amount of learning in creativity over time.

Unlike the two preceding MANOVAs, the pre-existing differences between OM groups were not found with the global multivariate test. It must be noted that in this analysis the uneven distribution of subjects per cell is even more marked than the other analyses, which also contained highly unequal cell sizes. In this analysis, three of the six cells contain fewer than 20 students ($\underline{n}$=10, 12, and 12 respectively). In the other 3x2x2 MANOVA for experience in the enrichment program, two cells contained fewer than 20 subjects with only 11 subjects in those two cells. The analysis contains only one case per dependent variable in the cell for OM-hi, no prior experience on an OM team. This analysis, with the large number of dependent variables and low sample size in some of the cells, therefore, is likely to reflect a loss of power. Appearance of several significant univariate effects without the accompanying multivariate effect could be an indicator of loss of power in the analysis (Tabachnick & Fidell, 1983). Examination of Table P-7, Appendix P, reveals five univariate Fs significant at $p < .05$ without the global effect.

A significant multivariate difference was found when examining the effect of prior experience on an OM team, $F(10,101)=4.76$, $p < .001$, $N^2=.32$, accounting for a large portion of the variance in the analysis. Support for the construct occurred in these univariate analyses (all df=1,110): Similes, $F=22.06$, $p < .001$, $N^2=.12$; total of creative strengths, $F=21.45$, $p < .001$, $N^2=.14$; closure, $F=13.47$, $p < .001$, $N^2=.06$; verbal flexibility, $F=4.48$, $p < .05$, $N^2=.03$; figural originality, $F=4.58$, $p < .05$, $N^2=.02$; elaboration, $F=4.19$, $p < .05$, $N^2=.03$. The analysis also yielded significant stepdown results for four of the six measures: Similes, $F(1,104)=9.91$, $p < .01$; total of creative strengths, $F(1,105)=9.04$, $p < .01$; closure, $F(1,106)=8.93$, $p < .01$; figural originality, $F(1,109)=4.52$, $p < .05$. A comparison of the pretest means for the respective six measures shows, as in the preceding analysis of the effect of experience, with the exception of verbal flexibility, no prior experience=137.18, team experience=130.09; the students with experience on an OM team (team experience) showed greater strength in measured creative thinking ability than those with no prior experience (no experience) on a team. It should be noted that the cell for non-OM, no experience, contained an extremely high cell mean for all three verbal measures, thus indicating the possibility of containing a disproportionate number of highly creative students.

The cell means for the variables which showed higher pretest scores for students with OM experience were: Similes, no experience=37.19, team experience=51.13; total of creative strengths, no experience=8.48, team experience=10.07; closure, no experience=107.54, team experience=110.33; figural originality, no experience=116.89, team experience=123.75; figural elaboration, no experience=125.98, team experience=133.33. These results indicate that OM may be an effective method for training creativity.

Summary of Multivariate Analyses of the Creative Construct

The results of these analyses fail to support the primary hypothesis that there would be a change in the dependent construct over time that would be different among groups with respect to OM effort. Likewise, the evidence did not support hypotheses 3, 5, 7, or 9. There was, however, a strong effect of change over time across groups which would likely be attributable to treatment effects of the overall enrichment program. These results showed a rise in scores in Similes, figural fluency, and originality but a lowering of scores on figural elaboration and all three measures of the verbal TTCT. The decrease in scores on the verbal TTCT conflicts with the consensus of research findings that there is an increase in scores in response to creativity training programs (Feldhusen & Clinkenbeard, 1986; Torrance, 1972a), and particularly on verbal creativity (Rose & Lin, 1984).

The implications for these findings will be discussed in Chapter V, and in Appendix E.

There were pre-existing differences between OM groups on the creativity construct such that students who volunteered to be on an OM team and would exert above average effort in their work on the team (OM-hi) showed higher pre-treatment scores on Similes, closure, and total of creative strengths than those who were not on a team. Some were higher even than those who would be categorized in the OM low effort group. There were also pre-existing differences between students who were experienced in the enrichment program, and who were experienced in OM and those who had less experience in enrichment or no prior experience on a team. An interesting finding is that both experience factors showed the same pattern of results. The multivariate analyses examined the various ten dependent variables for comparable contributions to the overall multivariate effect. These showed that with the exception of verbal flexibility the students with greater experience in the program and/or experience in OM showed evidence of higher verbal creative thinking abilities in Similes and of figural abilities in originality, elaboration, closure, and total of creative strengths. Although the primary hypothesis for the OM x Time interaction was not supported, these results lend support to the hypothesis that the enrichment program and OM were effective sources of training creativity.

Need for Supplemental Analyses

The findings of significant effects of the experience factors occurring as pre-existing differences among students in the study indicated that those could well be error variables in the study. Grade level, school, and sex were additional uncontrolled variables in the study. A series of multiple regression analyses were performed to test for violation of the assumptions for multivariate analysis (see Appendix B). It was discovered that grade was a significant predictor of all verbal variables, closure, creative strengths, and elaboration; and sex was a significant predictor of figural fluency, originality, elaboration, and the total of creative strengths. The uncontrolled variable for difference in distribution for the 10 schools was simplified by recoding school for possible differences between the three enrichment teachers. It was found that differences between teachers was not a significant predictor (see Table P-8, Appendix P).

Uncontrolled variable - grade level. To assess the affect of grade level, a 3x3x2 MANOVA was performed in which the independent variables were OM (non-OM, OM-lo, OM-hi), grade (2+3, 4, 5) and time (pre, post). Grades two and three were combined for this analysis to provide a better balance of numbers in the various OM and grade level cells, see Table F-3, Appendix F). The cell means and standard deviations are shown in Table P-9 in Appendix P. The summary of the MANOVA of grade and OM effort, in Table P-10 in Appendix P, shows a significant multivariate difference in creativity in the interaction of grade level and time $F(20,198)=9.67$, $p < .001$, $N^2=.76$. Support for the effect on the construct was found in these univariate analyses (df = 2,107): closure, $F = 31.68$, $N^2=.18$; verbal fluency, $F = 10.19$, $N^2=.04$; figural elaboration, $F = 7.58$, $N^2=.03$; all at $p < .001$; verbal originality, $F = 4.83$, $N^2=.02$, $p < .01$; and figural fluency, $F = 3.35$, $N^2=.02$, $p < .05$. In the examination of the effect of grade level and time on creativity, the stepdown analyses showed that closure $F(2,103)=27.29$, and verbal originality $F(2,98)=24.55$ had the greatest effect in the construct followed by total of creative strengths $F(2,102)=7.58$; figural elaboration $F(2,104)=7.16$; all at $p < .001$; verbal fluency $F(2,100)=6.93$, $p < .01$; and verbal flexibility $F(2,99)=4.08$, and figural fluency $F(2,107)=3.35$, $p < .05$, in this order.

The significant univariate interactions of Grade x Time (Figure P-11, Appendix P) showed that posttest scores were markedly higher than pretest for figural closure at Grade 5, but were somewhat lower than pretest at Grades 2+3 and 4. Verbal fluency pretest scores were remarkably high for Grades 2+3, with a mean standard score of 141.62. Posttest scores were markedly lower (126.17) for Grade 2+3 and barely different from the pretest scores at Grades 4 and 5. Figural elaboration pretest scores were likewise very high at Grades 2+3 and 4, and posttest scores were much lower at these grades with no difference at Grade 5. Verbal originality scores were considerably lower at posttest than for pretest for all grades, but less so for Grade 5. Posttest total of creative strengths scores were higher than pretest at fourth and fifth grade. Further examination of the cell means (see Table P-9, Appendix P) reveals that the greatest portion of the decline in verbal TTCT performance from pre- to posttest occurred for the following: verbal fluency, Grade 2+3 (pre=141.52, post=127.70), and verbal originality, Grade 2+3 (pre=141.12, post=120.93); verbal fluency, Grade 5 (pre=116.55, post=114.72), and verbal originality, Grade 4 (pre=121.45, post=101.84) and Grade 5 (pre=114.93, post=107.66). The remarkably higher pretest scores for Grade 2+3 when compared to Grade 4 may in part reflect the decline in creativity known as the "fourth grade slump" (Torrance, 1974a, 1974b; Williams, 1976; see Chapter V for discussion). Examination of cell means from a supplementary analysis for Grade 2 and for Grade 3 revealed that the decline for closure and elaboration occurred for Grade 3 but not for Grade 2: closure, Grade 2 (pre=97.77, post=103.50), Grade 3 (pre=113.93, post=105.18); elaboration, Grade 2 (pre=112.50, post=111.00), Grade 3 (pre=135.24, post=121.72). However, the decline in posttest scores occurred for both grades on the verbal TTCT measures. The decline in verbal scores in this study may also in large part be attributed to a "context effect" (see Chapter V and Appendix E).

In order to more clearly test the hypotheses of the study, it was deemed necessary to statistically control the significant pre-treatment variables which were shown to be predictors of posttest creative performance, grade and sex, by means of multivariate analysis of covariance (MANCOVA). The unequal distribution of the experience, grade, and sex factors as pre-treatment differences suggest that the more appropriate analysis would be MANCOVA.

Possible violation of assumptions of multicollinearity. The values on the intercorrelation matrix among the three verbal TTCT scores showed relationships that are excessively high for effective multivariate anlysis: fluency and originality r=.95; fluency and flexibility r=.91; and flexibility and originality r=.92; see Table P-1, Appendix P. Such inordinately high relationships between variables indicates possible problems with multicollinearity of the multivariate matrix. As discussed in Appendix B, the threat of multicollinearity was judged to be acceptably solved. When there are a large number of dependent variables in relation to the sample size as well as excessively high relationships among some of the dependent variables, combination of variables is recommended (Bray & Maxwell, 1985; Tabachnick & Fidell, 1983). Therefore, it was decided to combine variables by calculating the mean of the three verbal TTCT standard scores. A Creativity Index, the mean of the five standard scores, plus the raw score of total indicators of creative strengths, was calculated for the figural TTCT measures as suggested by Torrance and Ball (1984). Torrance and Ball reported a study which suggested evidence of predictive validity of each of the standardized scores of the streamlined scoring, the total indicators of creative strengths, and also of the Creativity Index. Chase (1985) suggested that a single score should be substituted for the three TTCT verbal measures because of excessively high intercorrelation.

Multivariate Analyses of the Abbreviated Creative Construct

The three dependent variables forming the abbreviated creative construct were Similes, the mean of the three TTCT verbal scores, and the figural Creativity Index, comprised of the mean of the five standard figural TTCT scores plus the raw score for total of creative strengths. The figural Creativity Index was given the highest priority as most remote from the independent variable in an a priori hierarchy of importance among the dependent variables. The remaining dependent variables were ordered such that Similes and then the mean verbal score would be adjusted by the covariate of the preceding dependent variables. Therefore, the mean verbal score, as the dependent variable of most critical interest, would receive a pure analysis.

Examination of the within cells correlation matrices of the dependent variables showed that a multivariate construct was formed only for the first analysis, the 3x2 mixed model MANOVA. In Table P-12, Appendix P, showing the respective intercorrelations of the three creative variables for the abbreviated construct, it is seen that Similes and the figural Creativity Index bear a closer relationship to each other than either measure bears to the mean verbal measure.

Effect of Odyssey of the Mind Effort

A 3x2 mixed model MANOVA was performed on the three creative dependent variables in order to test hypothesis 1, the OM x Time interaction. The fixed independent variables were OM (non-OM, OM-lo, OM-hi) for the between groups comparison and time (pre, post) for the repeated measure. Cell means and standard deviations are reported in Table 3.

For this analysis a multivariate construct was formed. The results of the MANOVA for OM effort showed a significant difference on the multivariate creative construct in pretest scores for OM groups $F(6,222)=4.33$, $p < .001$, Time, $F(3,111)=26.99$, $p < .001$, and OM x Time, $F(6,222)=2.20$, $p < .05$ as shown in Table 4. The results reflected a strong association between the effects of time (pre- to posttest scores) and the combined dependent variables, $N^2=.42$. The relation of OM, $N^2=.20$, and OM x Time, $N^2=.11$, to the construct were less marked. The significant findings for pre-existing differences between OM groups

Table 3

Cell Means and Standard Deviations of Three Creative Dependent Variables, Adjusted Creative Dependent Variables, and Odyssey of the Mind (OM) Effort[a]

		Dependent Variables					
		Similes		Mean Verbal TTCT		Figural TTCT Creativity Index	
OM Effort	n	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Pretest Means							
Non-OM	51	37.92	13.64	131.01	20.75	119.39	12.85
OM-lo	32	46.53	11.27	127.44	17.40	125.98	9.51
OM-hi	33	50.24	21.21	126.88	17.56	127.32	11.80
Posttest Means							
Non-OM	51	47.67	16.16	120.39	15.82	124.28	10.25
OM-lo	32	49.44	14.41	117.58	17.92	128.91	12.32
OM-hi	33	61.12	18.21	126.47	18.74	132.29	11.78
Adjusted Posttest Means[b]							
Non-OM		52.34		118.95		127.65	
OM-lo		48.09		118.04		127.97	
OM-hi		57.80		127.45		129.85	

Note. The verbal and figural TTCT scores are expressed in standard scores. Similes is expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at $p \leq 160$.
[b]Posttest scores adjusted for five covariates: preexisting differences between subjects in pretest scores, grade, and sex.

Table 4

Summary of Mixed Model Creative MANOVA for Odyssey of the Mind (OM) Effort: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Three Creative Dependent Variables[a]

Source	Multivariate Global Fs	Stepdown Fs/Univariate Fs: Dependent Variables: Similes	Mean Verbal TTCT	Figural TTCT Creativity Index
Between Subjects Analyses				
OM	4.33***[b]	4.55**[c] 8.84***	.94 .62	7.60*** 7.60***
Repeated Subjects Analyses				
Time	26.99***	33.87*** 28.92***	20.04*** 21.74***	14.12*** 14.12***
OM x Time	2.20*	2.72 2.34	3.56* 3.88*	.29 .29

[a]All standard scores set at $\leq$ 160.
[b]Wilks' Lambda criterion for statistical inference.
[c]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001.

and across groups for time were reported for the first analysis of the effect of OM effort and time on the ten creative dependent variables (see p. 146).

As the interaction between OM effort and time yielded a significant difference for the global multivariate construct, it is appropriate to consider the significant univariate F statistics associated with the three dependent variables for these effects. These statistics are reported in Table 4. Support for the construct of a difference in creativity scores between OM groups and time was found for the TTCT mean verbal measure in the univariate analysis (df=2,113), $F=3.88$, $p < .05$, $N^2=.02$; and in the stepdown analysis (df=2,111), $F=3.56$, $p < .05$. Post hoc comparisons using simple contrasts reveal that the mean verbal TTCT scores were higher for OM-hi than for OM-lo ($t=2.63$, $p < .01$), and for OM-hi than for non-OM ($t=2.19$, $p < .05$). A comparison of the mean scores in Table 3 shows that scores on the TTCT mean verbal went down from pre- to posttesting for students in the non-OM and OM-lo groups, non-OM (pre=131.01, post=120.39), OM-lo (pre=127.44, post=117.38), but did not change from pre- to posttesting for students whose effort on an OM team was ranked above the median, OM-hi (pre=126.88, post=126.47). The results are shown graphically in Figure 3.

The primary hypothesis that there would be a difference between creativity test scores among the different OM effort groups in change over time was, therefore, supported by the results of this analysis.

<u>Effect of Experience in the</u>

<u>Enrichment Program</u>

A 3x2x2 mixed model MANOVA was performed on the abbreviated creative construct to test hypothesis 3, the effect of the interaction of the amount of experience in the enrichment program, and hypothesis 7, the interaction of OM x Experience in Enrichment Program x Time. The intercorrelations among the three variables were less than .30 (see Table P-12, Appendix P), thus requiring a

Mean Verbal TTCT

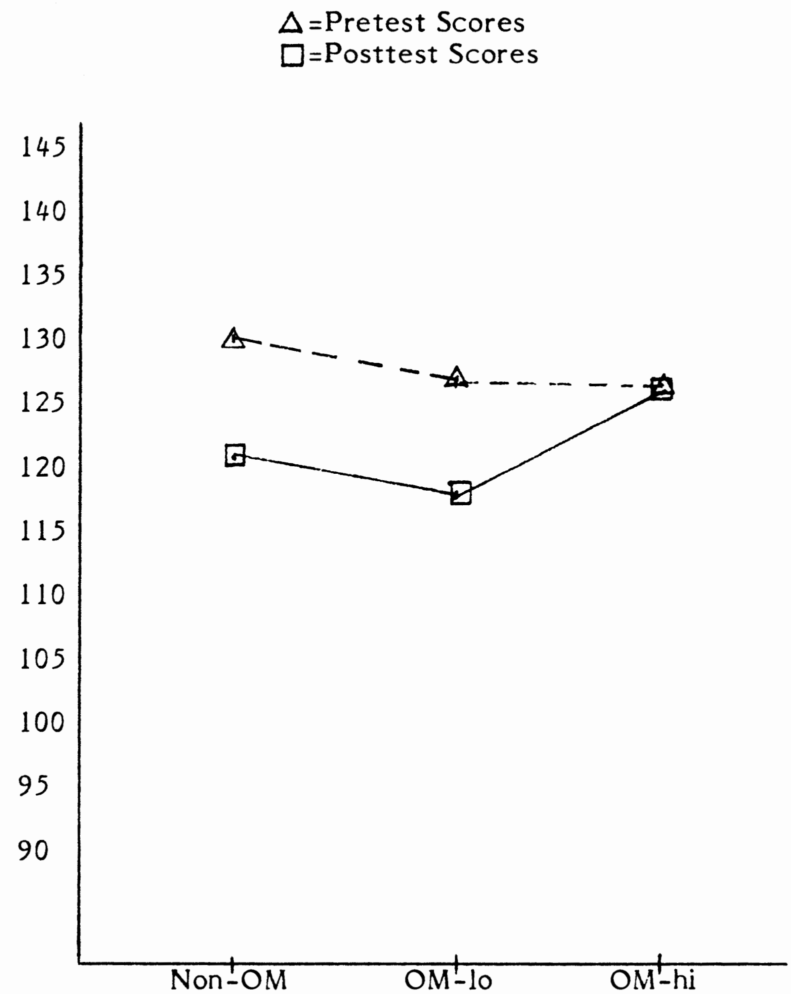


Figure 3. Odyssey of the Mind (OM) X Time Interaction for Creativity from Mixed Model MANOVA

Note. All standard scores set at ≤ 160.

univariate rather than multivariate interpretation. The independent variables were effort in OM (non-OM, OM-lo, OM-hi) and experience in the enrichment program (less than 13 months experience, 13 or more months experience) to compare effects between groups, and time (pre, post) to test for changes within individuals. Observed cell means and standard deviations are presented in Table P-13, Appendix P).

A summary of this 3x2x2 MANOVA is reported in Table P-14, Appendix P. As already reported in the first analysis of the abbreviated creative variables, a significant main effect was found for pre-existing differences between OM groups, for change in scores over time across groups, and for the interaction between OM and time. As in the analysis for all ten creative dependent variables, the interaction between experience in the program and time was not significant, thus hypothesis 3 was not supported. The three-way interaction was also not significant thus it did not provide support for hypothesis 7.

A significant pre-existing difference was found for experience in the enrichment program for two of the three variables: Similes, $F(1,110)=6.50$, $p < .001$, $N^2=.05$; and the figural Creativity Index, $F(1,110)=11.83$, $p < .001$, $N^2=.03$. A comparison of the means for these respective measures reveals that the students with 13 or more months of experience in the enrichment program showed greater evidence of cognitive creativity than students with less experience in the program: Similes, less than 13 months=38.65, greater than or equal to 13 months=46.52; figural Creativity Index, less than 13 months=119.87, greater than or equal to 13 months=125.35. These results might lend strength to previous findings cited in the literature that creativity can be enhanced by creativity training. The significant interaction of OM effort and time supports the findings that the primary hypothesis should be accepted, that is, a difference was found between creativity test scores among the different OM groups over time.

Prior Experience in Odyssey of the Mind

A 3x2x2 mixed model MANOVA was performed on the abbreviated creative construct to test hypothesis 5, the effect of the interaction of prior experience on an OM team and time, and hypothesis 9, the interaction of OM effort x Prior OM x Time. As in the preceding analysis for experience on the abbreviated measures, a multivariate construct of correlations $>$ .30 was not established (see Table P-12, Appendix P), thus requiring a univariate interpretation of the data. The independent variables were effort in OM (non-OM, OM-lo, OM-hi) and prior experience on an OM team (no prior experience, prior experience on an OM team), and time (pre, post). Cell means and standard deviations are presented in Table P-15, Appendix P.

A summary of this 3x2x2 MANOVA is reported in Table P-16 of Appendix P. As previously reported in the preceding analyses of OM effort and the creative construct, significant differences were found for pre-existing differences in OM on Similes and the Creativity Index, across groups over time for all three dependent variables, and on the interaction of OM and time for the TTCT mean verbal variable. As in the analysis of the ten creative dependent variables, the interaction between prior experience in OM and time was not found to be significant, therefore, not supporting hypothesis 5, nor the interaction between Prior OM x OM x Time, therefore, not supporting hypothesis 9.

A significant difference was found when examining pretest scores for the effect of prior experience on an OM team (all df=1,110): Similes, $F=22.06$, $p < .001$, $N^2=.12$; figural Creativity Index, $F=13.98$, $p < .001$, $N^2=.04$; and TTCT mean verbal, $F=3.96$, $p < .05$, $N^2=.03$. A comparison of the means for the respective measures shows, as in the preceding analysis of the effect of experience, that the students with experience on an OM team (team experience) showed greater strength in Similes and in the figural creative thinking abilities than those with no prior experience (no experience) on a team. However, those with team experience show lower TTCT verbal creativity than those with no experience. It should be noted that the cell for non-OM, no prior experience on a team contained a high cell mean and high standard deviation for the mean verbal measure, thus indicating the possibility of containing a disproportionate number of highly creative students.

The cell means for the variables which showed higher pretest scores for students with OM experience were: Similes, no experience=37.19, team experience=51.13; figural Creativity Index, no experience=120.98, team experience=126.53; and TTCT mean verbal, no experience=132.61, and team experience=124.68.

The lower verbal TTCT creativity in children with team experience appears to conflict with the finding that children who are high in effort in OM showed no decrease in verbal creativity over time. These findings will be further discussed in Chapter V. The results for Similes and the Creativity Index, however, indicate that experience in OM enhances children's creativity. The significant interaction of OM and time found in this analysis again lends support for the primary hypothesis that there would be a difference over time in creativity test scores between children in the different OM groups.

Need for Analysis to Control Pre-Existing Differences Among Subjects

The preceding analyses show a strong effect of change over time in measured creativity. Also shown in the preceding analyses is that there were pre-existing differences within the groups which showed as significant effects on pretest creativity scores. The pre-existing differences in amount of previous experience in the enrichment program and prior experience on an OM team (see Tables F-1 and F-2, Appendix F), probably serve to increase the error variance and consequently confuse possible effects of the major independent variable of interest, that is, effort on an OM team. Moreover, there was an uneven distribution in OM groups of subjects with regard to grade level and sex (see Table F-3, Appendix F). Grade level was shown to account for significant differences on pretest scores in creativity, as well as to account for some of the variance in changes over time (see Table P-10, Appendix P). Sex was shown by the regression analyses to significantly contribute to the prediction equation of several of the dependent variables (see Table P-8, Appendix P). These four variables impact the measured creativity of students in the sample and are all unequally distributed in the sample with respect to the independent variable of primary interest, OM effort. It was, therefore, judged necessary to use a multivariate analysis of covariance (MANCOVA) to control the effects of the pre-existing differences in order to examine the effects of OM effort more clearly.

Experience in enrichment and in OM, although shown to account for significant pre-existing differences in student creativity, however, were not significant predictors of posttest performance for any of the dependent variables according to regression analyses of the ten posttest dependent variables (see Table P-8, Appendix P) and the three posttest dependent variables (Table P-17, Appendix P). "Useless covariates" (Tabachnick & Fidell, 1983, p. 204), that is, ones which do not significantly improve upon the prediction of the regression equation, should be eliminated as covariates from analysis of covariance because each additional covariate reduces power by taking up degrees of freedom. Therefore, as recommended by Tabachnick and Fidell, preliminary MANCOVAs were performed which confirmed that the optimum set of covariates for the analysis of covariance of the posttest creativity data consisted of the pretest scores, grade and sex, and the prediction equation improved when the two potential experience covariates were eliminated from the analysis.

Multivariate Analysis of Covariance of the Abbreviated Creative Data

A one-way between subjects multivariate analysis of covariance (MANCOVA) was performed on the posttest scores of the abbreviated creativity construct: Similes, TTCT mean verbal, and the figural TTCT Creativity Index. The purpose of this analysis was to test hypothesis 1, the effect of the interaction of OM effort and time, in an indirect manner by testing for the effects of differences in OM effort groups on adjusted posttest scores. The abbreviated set of dependent variables was chosen for the MANCOVA over the full set of ten dependent variables to prevent loss of power by using too many degrees of freedom for the covariates. Use of multiple covariates, particularly when they are highly correlated with one another, could negate the gain in power due to controlling the error variance through covariance (Tabachnick & Fidell, 1983). Adjustment was made for five covariates: the three respective pretest scores, grade level, and sex. No outliers were identified by means of $SPSS^{X}$ (SPSS, 1986) regression analyses for each variable at $p < .05$. Results of evaluation of assumptions of normality, homogeneity of covariance matrices, linearity, multicollinearity, and homogeneity of regression were judged to be satisfactory (see discussion in multivariate procedural issues: tests of the assumptions of MANOVA and MANCOVA, Appendix B). Covariates were judged to be adequately reliable for covariance analysis.

The combined dependent variables were judged by Wilks' criterion to be significantly related to the combined covariates, approximate $F(15,293)=9.16$, $p < .001$ (see Table 5). The combined covariates (all at $df=5,108$) were significantly related at $p < .001$ to each of the dependent variables: Similes

$F=9.05$, TTCT mean verbal, $F=11.30$, and figural Creativity Index, $F=8.53$. The results indicated a strong association between the creative construct and the covariates with $N^2=.65$.

The power of the covariates to adjust the dependent variables was investigated by multiple regression of each dependent variable with the covariates as multiple predictors. For each dependent variable the pretest provided a significant prediction of its corresponding posttest scores. Pretest scores on these abbreviated dependent variables each bore a strong predictive relationship to their respective posttest scores with β values greater than .34, significantly different from zero, $t \geq 3.98$, $p < .001$. Additional significant predictors to the overall multivariate multiple regression were pretest scores on the figural Creativity Index as a predictor of posttest scores on Similes, $\beta=.27$, $t=3.16$, $p < .01$; and pretest scores on Similes on the posttest Creativity Index, $\beta=.18$, $t=2.07$, $p < .05$. Only one of the two covariates which were added to control excessive variability among subjects in this quasi-experimental study was a significant predictor of the three dependent variables. Difference in grade was an important predictor of posttest figural Creativity Index scores, $\beta=.28$, $t=2.68$, $p < .01$; and TTCT mean verbal scores, $\beta=-.24$, $t=-2.34$, $p < .05$. The covariate of sex did not provide significant adjustment to any of the dependent variables. The cell means and standard deviations and adjusted cell means are reported in Table 3. The effects of OM on the adjusted posttest scores were tested on the one-way MANCOVA of the adjusted creative construct. These results are presented in Table 5.

The intercorrelation matrix showed a very weak relation between the adjusted dependent variables (see Table P-12, Appendix P). The multivariate construct was not established, thus requiring univariate interpretation of this analysis.

Table 5

Summary of Creative MANCOVA for Odyssey of the Mind (OM) Effort: F of Multiple Regression and Univariate Fs for Three Creative Dependent Variables[a]

Source	Multiple Regression F	Univariate Fs: Dependent Variables: Similes	Mean Verbal TTCT	Figural TTCT Creativity Index
Within Cells Regression Analysis of Covariates				
Regression F	9.16**[b]			
Multiple R		.54	.59	.53
F-tests for variables		9.05	11.30	8.53
MANCOVA Between Subjects Analysis				
OM		3.84*	4.10*	.46

Note. Posttest scores adjusted for five covariates: preexisting differences between subjects in pretest scores, grade, and sex.

[a]All standard scores set at $\leq$ 160.
[b]Wilk's Lambda criterion for statistical inference.

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

After statistically adjusting for differences in grade, sex, and in pretest scores, significant differences in adjusted posttest scores showed on two of the three measures for the different OM groups. Adjusted posttest scores were significantly different for Similes, $F(2,108)=3.84$, $p < .05$, $N^2=.07$ and mean verbal TTCT, $F(2,108)=4.10$, $p < .05$, $N^2=.08$. Post hoc comparisons using simple contrasts and $p < .05$ revealed that, after adjustment of posttest scores by pretest scores and variability on the other two covariates, Similes scores were significantly higher for the OM-hi group than OM-lo, $t=-2.77$, $p < .01$. Adjusted mean verbal scores were significantly higher for the OM-hi group than for OM-lo ($t=-2.63$, $p < .05$) and higher than for non-OM ($t=-2.36$, $p < .05$). These differences are shown graphically in Figure 4.

It would seem, therefore, that when examining the data by controlling the important uncontrolled variables, amount of effort on an OM team can indeed be interpreted to correspond to differences in measured creativity. Students who exerted high effort on an OM team were higher in verbal creativity on verbal TTCT and Similes than students who either exerted low effort in OM or were not on an OM team. The OM-lo effort students did not show as high scores on Similes as those in either of the other two groups. These results offer clear support for hypothesis 1.

Tests of the Effects of the Affective Construct

Multivariate Repeated Analyses of the Affective Data

The four dependent variables forming the affective construct were general self-concept, Sears; locus of control for responsibility (two scores), I+, I-; and creative self-concept, Ideas. Sears self-concept was given the highest priority as

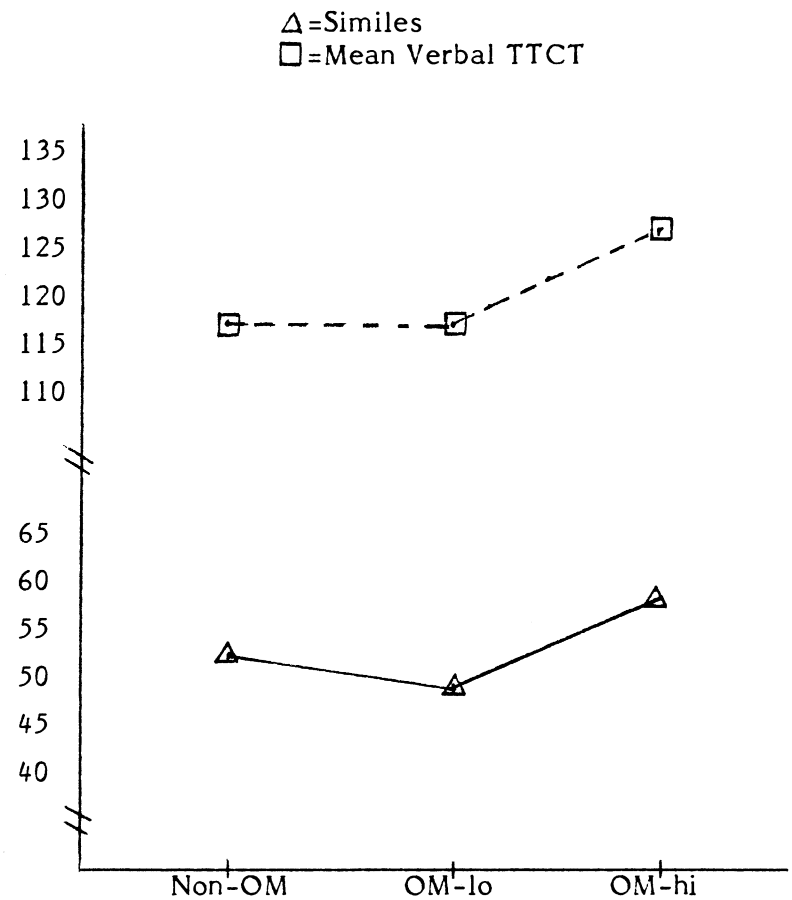


Figure 4. Effect of Odyssey of the Mind (OM) on Similes and on Mean Verbal TTCT Posttest Scores Adjusted by Five Covariates in MANCOVA

Note. The verbal and figural TTCT scores are expressed in standard scores. Similes is expressed in raw scores. See Chapter III, Instrumentation.
Note. All standard scores set at ≤ 160.

being the most remote from the independent variable in an a priori hierarchy of importance among the dependent variables. The remaining dependent variables were ordered such that I+, then I-, and lastly Ideas would be adjusted by the covariate of the preceding dependent variable(s) as well as the four covariates. Therefore, Ideas, as the dependent variable of most critical interest, would receive a pure analysis.

Examination of the within cells correlation matrices of the dependent variables for affect indicated that a multivariate construct was formed for all analyses (see Table P-18, Appendix P). Creative self-concept was more highly correlated with overall self-concept (pretest $r=.64$) than it was with the locus of control measures ($r_{I+}=.21$, $r_{I-}=-.28$).

Effect of Odyssey of the Mind Effort

A 3x2 mixed model MANOVA was performed on the affective construct consisting of the four dependent variables. The primary purpose of this analysis was to test hypothesis 2, the major hypothesis of the study, the OM x Time interaction. The secondary purpose of this analysis was to test hypotheses 11 and 12, that there would be differences among the univariate variables of the multivariate construct with respect to OM, and in creative self-concept with respect to OM x Time. The independent variables were effort in OM (non-OM, OM-lo, OM-hi) to compare effects between groups, and time (pretest, posttest) to compare changes within individuals. Observed cell means and standard deviations for each of the four affective measures are presented in Table 6.

In Table 7 the multivariate analyses for OM effort testing the major hypothesis of the study, the interaction between OM effort and time, shows that the global multivariate construct did not yield significance, $F(8,210)= 1.42$, $p < .05$. However, as for the creative construct, significant main effects were

Table 6

Cell Means and Standard Deviations of Affective Dependent Variables, Adjusted Affective Dependent Variables, and Odyssey of the Mind Effort

		Dependent Variables							
		Sears		I+		I-		Ideas	
OM Effort	n	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Pretest Means									
Non-OM	49	181.86	32.12	13.67	2.27	10.00	3.65	33.16	6.60
OM-lo	29	181.03	30.49	13.79	2.19	9.07	2.45	35.03	6.86
OM-hi	33	185.36	30.88	13.39	2.45	10.64	3.01	35.85	6.24
Posttest Means									
Non-OM	49	182.14	30.17	14.00	2.35	10.35	3.25	35.61	6.17
OM-lo	29	169.21	33.84	13.31	3.08	10.10	2.87	34.62	5.81
OM-hi	33	182.21	36.68	14.33	2.18	11.42	2.41	38.21	6.45
Adjusted Posttest Means									
Non-OM	49	185.78		13.73		10.60		35.99	
OM-lo	29	168.75		13.33		10.52		34.40	
OM-hi	33	179.03		14.59		10.77		38.05	

Note. Posttest scores adjusted for seven covariates: pre-existing differences between subjects in pretest scores, months of experience in enrichment program, sex, and years of prior experience in OM.

Table 7

Summary of Mixed Model Affective MANOVA for Odyssey of the Mind (OM) Effort: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Four Affective Dependent Variables

		Stepdown Fs/Univariate Fs			
		Dependent Variables			
Source	Multivariate Global Fs	Sears	I+	I-	Ideas
Between Subjects Analyses					
OM	1.99*[a]	.73[b] .73	.03 .21	2.69 2.31	4.51* 2.51
Repeated Subjects Analyses					
Time	4.81***	2.77 2.76	2.34 1.81	3.77 5.49*	9.49** 6.82**
OM x Time	1.42	2.21 2.21	2.47 2.88	.42 .53	.66 1.88

[a]Wilks' Lambda criterion for statistical inference.
[b]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001

found for pre-existing differences in OM between groups and for time across OM groups.

A significant multivariate difference was found on pretest scores between OM groups, $F(8,210)=1.99$, $p < .05$, $N^2=.14$. This multivariate effect was not supported by the univariate analyses, but was supported by the stepdown analyses for Ideas $F(2,105)=4.51$, $p < .05$. Post hoc comparisons using simple contrasts and $p < .05$ showed the OM-hi group to hold higher creative self-concept than the non-OM group ($t=-2.17$), see Figure 5. Cell means for Ideas show higher pretest scores for those choosing to be on a team, OM-hi=35.85, OM-lo=35.03, than those who did not participate in OM, non-OM=33.16. Prior to the treatment, students who elected to be in OM held significantly higher creative self-concepts than students who were not in OM.

When examining time, a significant multivariate difference was found for affect $F(4,105)=4.81$, $p < .001$, $N^2=.15$. Support for the construct from the univariate Fs was found in creative self-concept $F(1,108)=6.82$, $p < .01$, $N^2=.02$ and in attribution for negative events (I-), $F(1,108)=5.49$, $p < .05$, $N^2=.01$. The analyses also resulted in a significant stepdown $F(1,105)=9.49$, $p < .01$ for creative self-concept. Inspection of the means shows an overall increase in scores from pre- to posttest for both measures: I- (pre=9.95, post=10.60), Ideas (pre=34.45, post=36.12), see Figure 6.

These results show that, based on pretest scores, the OM effort groups were not equivalent in creative self-concept, and there was a significant effect on student affect over time such that there was an overall increase in locus of control for negative events and in creative self-concept. The primary hypothesis that there would be a difference between affective test scores among the different OM effort groups in change over time was not supported.

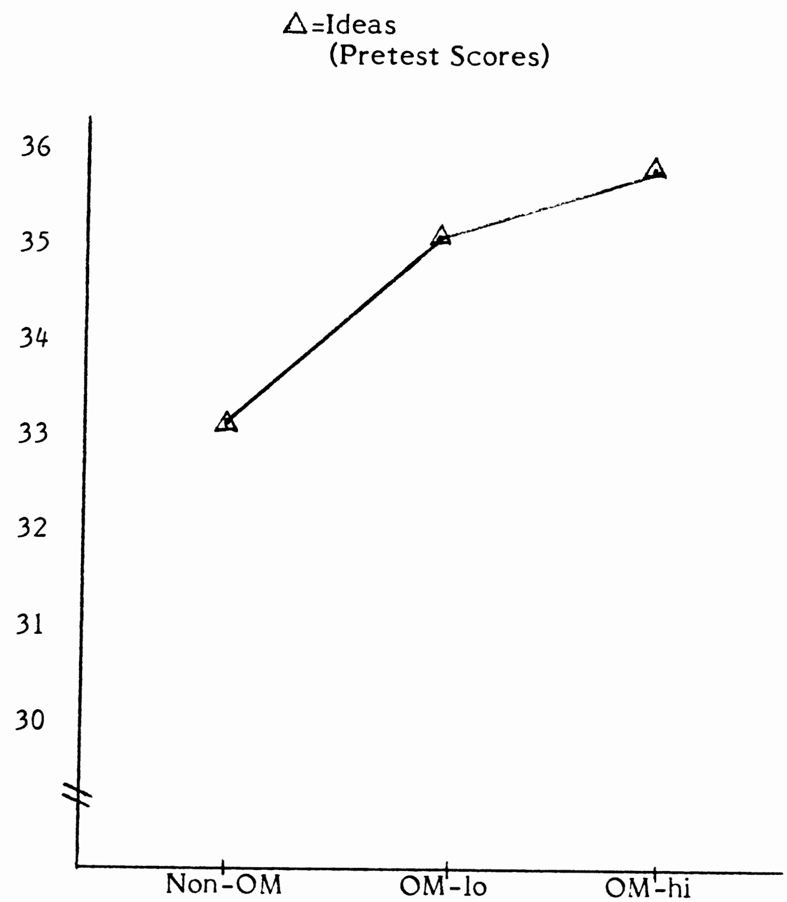


Figure 5. Preexisting Differences Between Odyssey of the Mind (OM) Effort Groups in Creative Self-Concept (Ideas)

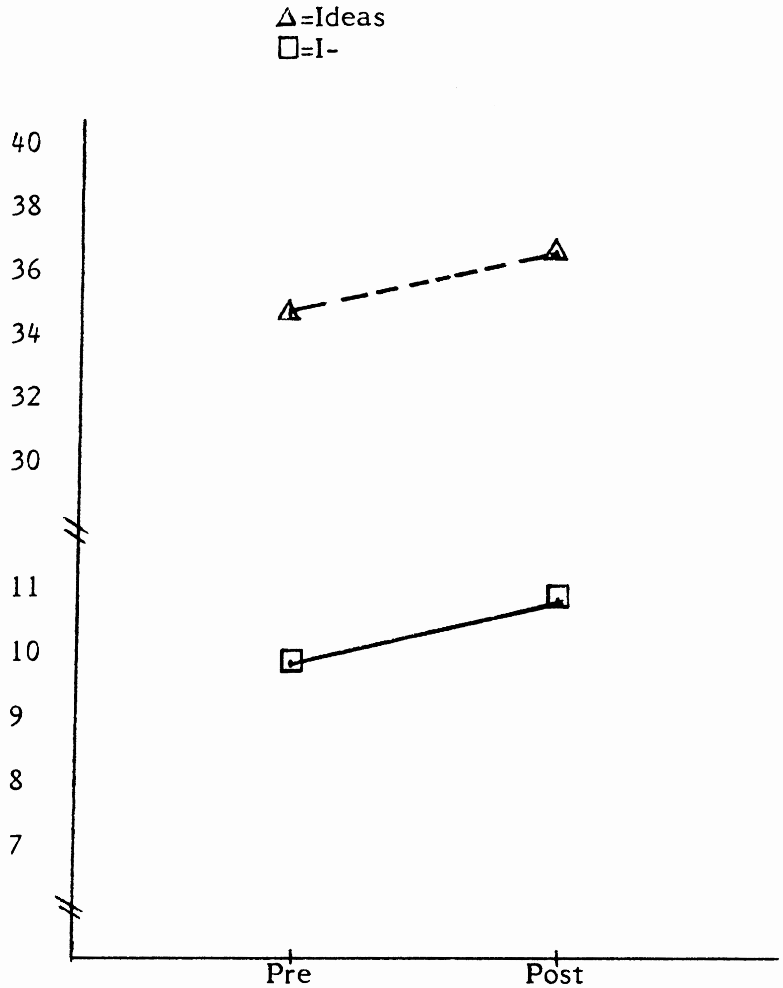


Figure 6. Main Effect of Time on Affective Construct Mixed Model MANOVA

Hypothesis 11, that there would be differences in sensitivity of the scores of the univariate dependent variables with respect to OM, was supported by the significant stepdown finding of differences in pretest scores on creative self-concept. Hypothesis 12, that there would be differences in creative self-concept scores with respect to student effort in OM and time, however, was not supported by the interaction analysis. The difference in sensitivity of these dependent variables over time also lends support to hypotheses 11 and 12. This indicates that responsibility for attribution toward negative events, as well as creative self-concept, were more sensitive to the effects of time, and, perhaps, also to the creativity training effects of the overall enrichment program, than were the remaining variables of the affective construct.

Effects of Experience in the Enrichment Program

A 3x2x2 mixed model MANOVA was performed on the affective construct to test hypothesis 4, the effect of the interaction of the amount of experience in the enrichment program and time, and hypothesis 8, the interaction of OM x Experience in Enrichment Program x Time. The independent variables were OM (non-OM, OM-lo, OM-hi) and experience in the enrichment program (less than 13 months experience, 13 or more months experience) to compare effects between groups, and time (pre, post) to test for changes within individuals. Observed cell means and standard deviations are reported in Table P-19, Appendix P.

A summary of this 3x2x2 MANOVA is reported in Table P-20 in Appendix P. As already reported in the first affective analysis, a significant main effect was found for pre-existing differences between OM groups and for significant changes over time across groups. No significant differences were found to be associated with experience in the enrichment program. The tests for hypothesis 4, the effect of experience in the enrichment program over time, and hypothesis 8, the interaction of OM x Experience x Time, failed to find a significant multivariate effect.

Prior Experience in Odyssey of the Mind

A 3x2x2 mixed model MANOVA was performed on the affective construct to test hypothesis 6, the effect of the interaction of prior experience on an OM team and time, and hypothesis 10, OM x Prior OM x Time. The independent variables were OM (non-OM, OM-lo, OM-hi) and prior experience on an OM team (no prior experience, prior experience on an OM team), and time (pre, post). Cell means and standard deviations are reported in Table P-21 in Appendix P.

A summary of the 3x2x2 MANOVA is reported in Table P-22 in Appendix P. As previously reported under the first analysis of OM effort and the affective construct, significant differences were found across groups over time. As with the factor of experience in the enrichment program, the multivariate analysis failed to find a significant global effect for prior experience on an OM team, but yielded a significant stepdown effect. As in the analysis for the creative construct (see p. 152), this MANOVA, unlike the two preceding affective analyses, did not find a difference between the OM groups.

The results of the affective analyses, then, do not support the hypotheses. No difference was found in change in affective scores over time with respect to effort in OM. There were no significant differences pre-existing in affect among OM groups, either of the experience factors, nor their interactions over time nor their interaction with the OM factor.

Need for Supplemental Analyses

The possible findings, however, that the experience factors might explain a pre-existing effect offer logical sense and are congruent with the significant effect over time across groups. The strong effect across groups over time suggests that there is a powerful influence to increase creative self-concept and attribution of responsibility for negative events as effects of the overall enrichment program. The overall strong effect found over time serves as a confirmation that the univariate finding of difference in attribution of responsibility in relation to the experience in the program is a real effect which is not showing its full significance due to low power in the 3x2x2 analyses.

The unequal distribution of the experience factors among the children in the sample, as well as the tendency for both experience factors to show as pre-treatment differences which did not change over time, suggests that the more appropriate analysis would be analysis of covariance. A MANCOVA was, therefore, performed to examine the effects of OM effort in a pure manner. The analysis of covariance would control for variations in individuals due to differences in length of experience in the enrichment program and/or in OM as well as control for sex as an uncontrolled variable. Grade level was determined to be a "useless covariate" (Tabachnick & Fidell, 1983) to be eliminated from the analysis of covariance for two reasons: test of regression slope (see Appendix B) found an interaction between OM and grade, and grade was not a significant contributor to the regression equation (see Table P-23, Appendix P).

Multivariate Analysis of Covariance of the Affective Data

A one-way between subjects MANCOVA was performed on the posttest scores of the four dependent variables of the affective construct: Sears, I+, I-, and Ideas. The purpose was to test hypothesis 2, the effect of the interaction of OM effort and time, in an indirect manner by testing for the effects of differences in OM effort groups on adjusted posttest scores. Adjustment was made for seven covariates: the four respective pretest scores, sex, months of experience in the enrichment program, and years of prior experience on OM team(s). No outliers were identified by means of $SPSS^X$ (SPSS, 1986) regression analyses for each variable at $p < .05$. Results of evaluation of assumptions of normality, homogeneity of covariance matrices, linearity, and multicollinearity were judged to be satisfactory (see discussion in multivariate procedural issues: tests of the assumptions of MANOVA and MANCOVA, Appendix B). Test of the assumption of homogeneity of regression showed that there is a problem with parallelism of the slopes for I- and prior experience in OM as covariates. Covariates were judged to be adequately reliable for covariance analysis.

The combined dependent variables were judged by Wilks' criterion to be significantly related to the combined covariates, approximate $F(28,354)=7.38$, $p < .001$ (see Table 8). The combined covariates (all at df=7,101) were significantly related at $p < .001$ to each of the dependent variables: Sears ($F=17.97$), I+ ($F=7.31$), I- ($F=7.25$), and Ideas ($F=6.31$). The results indicated a strong association between the affective construct and the covariates with $N^2=.80$.

The power of the covariates to adjust the dependent variables was investigated by multiple regression of each dependent variable with the covariates as multiple predictors. For each dependent variable the pretest provided a significant prediction of its corresponding posttest score. Pretest scores on the Sears, I+, and I- each bore a strong predictive relationship to their respective posttest scores with β values greater than .49, significantly different from zero, $t \geq 5.53$, $p < .001$, and for Ideas, $\beta=.27$, $t=2.67$, $p < .01$. Additional significant predictors to the overall multivariate multiple regression were pretest scores on the Sears as a predictor of posttest scores on I+, $\beta=.20$, $t=2.03$, $p < .05$ and on Ideas, $\beta=.25$, $t=2.50$, $p < .05$. All three covariates which were added to control excessive variability among subjects in this causal comparative study were significant predictors. Difference in sex was an important predictor of posttest Sears scores, $\beta=-.19$, $t=-2.83$, $p < .01$, and differences in years of OM on posttest

Table 8

Summary of Affective MANCOVA for Odyssey of the Mind (OM) Effort: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Four Affective Dependent Variables

Source	Multivariate Global F	Stepdown Fs/Univariate Fs			
		Dependent Variables			
		Sears	I+	I-	Ideas
Within Cell Regression Analysis of Covariates					
Regression F	7.38**[a]				
Multiple R		.74	.58	.58	.55
F-tests for variables		17.97**[b] 17.97**	5.63** 7.31**	6.73** 7.25**	1.86 6.31**
MANCOVA Between Subjects Analysis					
OM	2.35*	4.35* 4.35*	2.20 2.54	.12 .09	2.76 3.31*

Note. Posttest scores adjusted for seven covariates: preexisting differences between subjects in pretest scores, months of experience in enrichment program, sex, and years of prior experience in OM.

[a]Wilk's Lambda criterion for statistical inference.
[b]Stepdown Fs are above univariate Fs.

*$p < .05$. **$p < .001$.

I-, β=.31, t=2.37, $p < .05$. Months of experience in the enrichment program provided significant adjustment to posttest Ideas scores, β=.30, t=2.23, $p < .05$. The cell means and standard deviations and adjusted cell means and standard deviations are reported in Table 6.

The effects of OM on the adjusted posttest scores were tested on a one-way MANCOVA of the adjusted affective construct. These results are presented in Table 8. After statistically adjusting for differences in sex, experience in the program and in OM, and for pretest scores, a significant global effect was found for OM $F(8,196)=2.35$, $p < .05$. These results yield a weak association between the construct and the OM independent variable, N^2=.17.

Two dependent variables, with statistical adjustment for the seven covariates, made a significant contribution to the composite adjusted dependent variable that distinguished between members of the OM groups. Support for the global effect was found in the univariate analyses for Sears $F(2,101)=4.35$, $p < .05$, N^2=.09; and Ideas, $F(2,101)=3.31$, $p < .05$, N^2=.07. The analyses also resulted in a significant stepdown $F(2,101)=4.35$, $p < .05$ on adjusted Sears posttest scores. Inspection of cell means showed that Sears scores were highest in the non-OM group, and scores were markedly lower in the OM-lo effort group: Sears (non-OM=185.78, OM-lo=168.75, OM-hi=179.03). Post hoc comparisons using a simple contrast at $p < .05$ showed non-OM to be significantly higher in general self-concept than OM-lo, t=-2.95, $p < .01$ (see Figure 7). However, creative self-concept showed OM-hi to be higher than the OM-lo group (t=-2.57, $p < .05$). Creative self-concept was highest for the OM-hi effort group: Ideas (non-OM=35.99, OM-lo=34.40, OM-hi=38.05), see Figure 8. These results offer support for hypothesis 2, the primary hypothesis, that there would be a difference between OM effort groups in their adjusted posttest scores of the affective

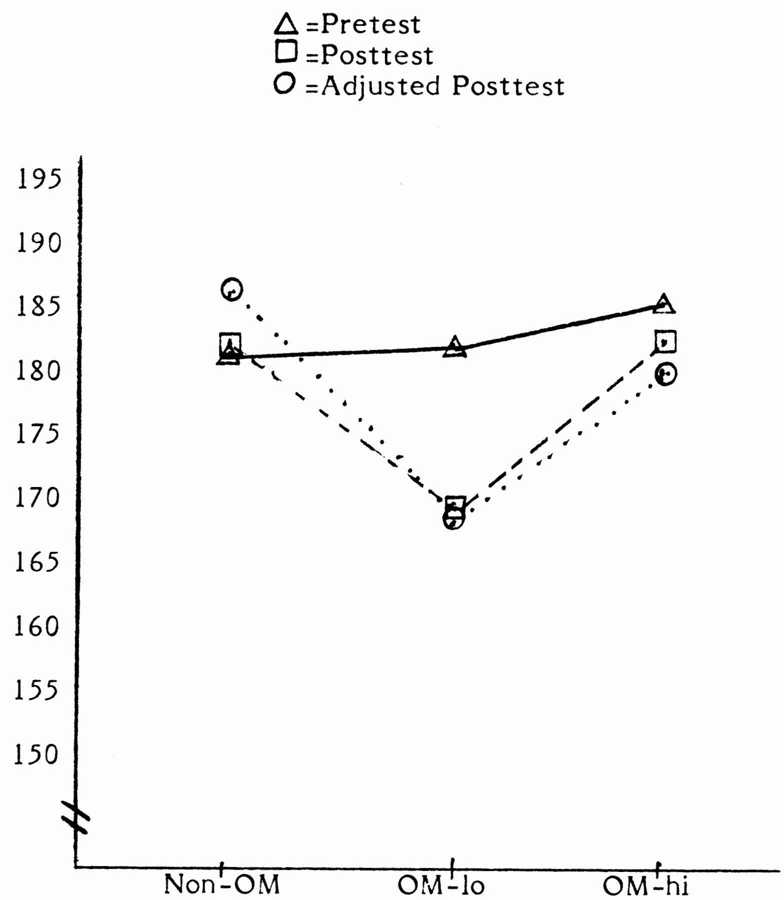
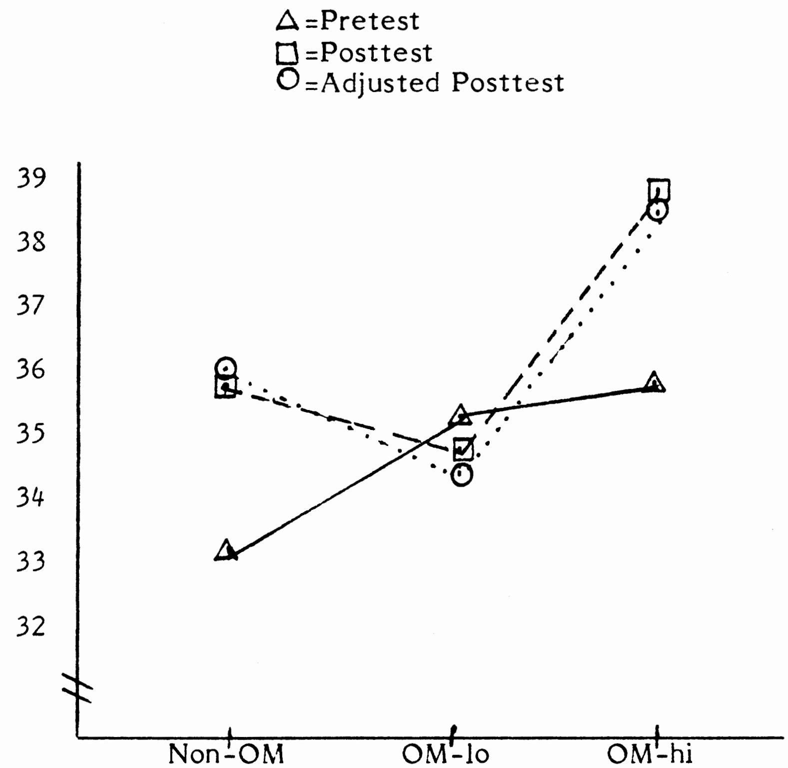


Figure 7. Effect of Odyssey of the Mind (OM) on General Self-Concept (Sears): Pretest and Posttest Sears Scores From Mixed Model MANOVA; Adjusted Posttest Scores from MANCOVA

Note. Posttest scores adjusted for seven covariates: pre-existing differences between subjects in pretest scores, months of experience in enrichment program, sex, and years of prior experience in OM.



<u>Figure 8</u>. Effect of Odyssey of the Mind (OM) on Creative Self-Concept (Ideas): Pretest and Posttest Ideas Scores From Mixed Model MANOVA; Adjusted Posttest Scores from MANCOVA

<u>Note</u>. Posttest scores adjusted for seven covariates: pre-existing differences between subjects in pretest scores, months of experience in enrichment program, sex, and years of prior experience in OM.

construct. Moreover, these results support hypothesis 12, that there is a difference in scores of creative self-concept in relation to student effort in OM.

Summary of Affective Analyses

The mixed model analysis of the multivariate construct for affect yielded a significant global effect across OM groups for time. Posttest scores were higher than pretest scores for creative self-concept and for locus of control for negative events. There was a pre-existing difference between OM groups such that children who elected to join an OM team were higher in creative self-concept than children who did not join a team. The affective analyses found no significant differences for any effects pertaining to either of the experience variables, experience in the enrichment program and prior experience in OM. Therefore, the findings failed to support the hypotheses that there would be differences on the interactions between either of these variables and time. The supplemental analysis by MANCOVA yielded a significant difference in adjusted posttest scores for OM, thus providing evidence supporting the primary hypothesis that there is a difference in OM groups over time and that creative self-concept differs with respect to OM groups over time.

It would seem, then, that students who were on an OM team and are judged by themselves and their peers to have put below average effort into the team's work perceive themselves with a less positive sense of self than do gifted students who were not on an OM team or who were on a team and were rated as expending above average effort. Students who were high effort showed a greater gain in creative self-concept than students who were in OM and experienced low effort in their work on an OM team. The creative self-concept scale showed sensitivity to changes in creativity training associated with the overall enrichment program and with OM.

Summary

In this chapter the two constructs, creativity and affect, were analyzed by three sets of mixed model MANOVA. These analyses tested the effect on the dependent variables of effort in OM and time, and also experience in the enrichment program or prior experience on an OM team. All analyses yielded a significant effect for differences in time, and most showed that pre-existing differences were present between the OM groups on scores of the dependent variables. The planned analyses, however, failed to support the major hypothesis that differences between groups of students on the dependent variables would occur over time in relation to their amount of effort on an OM team.

It was deemed necessary to adjust the posttest scores for the creative and affective dependent variables for the various important uncontrolled variables: differences in grade level, sex, and amount of experience in the program, and in OM. With examination of the adjusted variables through MANCOVA, the hypothesized relation became clear that differences in OM effort explained change in these gifted students in both creativity and affect. The new scale of creative self-concept showed sensitivity to the training effects associated with the overall enrichment program and with OM.

CHAPTER V

DISCUSSION

In this chapter the major findings are summarized and discussed in their relation to the twelve hypotheses of the study. Some alternative interpretations of the findings are offered, including effects of the varying statistical procedures. Assessment is made of those findings which shed light on the sensitivity and relevance of the instruments used in this study.

Suggested generalizability of the study and implications for future research are offered. The possible important contributions of this research are suggested.

Summary of Treatment Effects

Support for the Hypothesized Treatment Effects Over Time

Creativity Hypotheses

Each of the five hypotheses for the creative construct specified that a significant interaction would occur. The hypotheses were stated in a nondirectional manner for the following predicted effects:

1. There will be a significant interaction on the creativity construct score with respect to effort in OM and time.

3. There will be a significant interaction on the creativity construct scores of students in relation to the amount of their experience in the enrichment program and time.

5. There will be a significant interaction on the creativity construct

scores of students in relation to the amount of their prior experience in OM and time.

7. There will be a significant interaction on the creativity construct scores when these scores are compared with respect to effort in OM, experience in the enrichment program, and change in score from pretest to posttest in the experimental year.

9. There will be a significant interaction on the creativity construct scores when these scores are compared with respect to effort in OM, prior experience in OM, and change in score from pretest to posttest in the experimental year.

Of these five hypotheses, only one, the primary hypothesis for this study, was supported by the results. A difference in creativity scores was shown in change of scores over time in relation to student effort in OM. This OM x Time result occurred on the verbal TTCT creativity measures but not on the figural TTCT nor Similes for the mixed model analysis. The MANCOVA analysis, however, after adjusting posttest scores for pretest scores, and two uncontrolled variables, grade level and sex, yielded a significant effect for OM on Similes, as well as on the mean verbal TTCT scores (see Table 5 and Figure 4).

In contrast to the majority of results in the published literature on other effects of training for creativity (Cohn, 1985; Feldhusen & Clinkenbeard, 1986; Rose & Lin, 1984; Torrance, 1972a), there was a marked decline in verbal TTCT performance for two of the groups (non-OM and OM-lo) which experienced the overall creativity training of the enrichment program. The OM-hi effort group, however, was significantly different from the OM-lo and non-OM groups and showed no decline in verbal TTCT posttest performance. Examination of the pretest, posttest, and adjusted posttest cell means for Similes (Table 3 and Figure 4) shows that there was a marked increase in verbal creativity as measured by the Similes test for students in non-OM and OM-hi groups and a trend toward increase for OM-lo.

The difference between these effects of decline in scores for verbal TTCT but increase in scores for Similes is best explained by the detrimental effects on motivation noted during verbal TTCT posttesting. The sensitivity of the TTCT to motivational factors in the testing administration environment is discussed more fully below under main effects of time and in Appendix E (Elkind, Deblinger & Adler, 1970; Torrance, 1972a, 1987). A brief comment, then, will suffice at this point; it seems likely that there was a significant context effect as a detrimental effect on the children's motivation to show their potential creative energy because the students in this study were tested during a class which most experienced as considerably more enjoyable than the test activities. The reduction of scores was noted on all three verbal TTCT measures and on figural elaboration.

There were, then, higher posttest scores in Similes for the OM-hi group than for non-OM and OM-lo, and the mean verbal scores did not decline for the OM-hi group despite the likelihood of a significant deleterious context effect on the posttest verbal scores for children in this study. These results, therefore, show evidence to support the primary hypothesis of this study that there is a differential effect on posttest scores between the OM groups. More specifically, there was a decline in posttest verbal TTCT scores for the non-OM and OM-lo children, however, the OM-hi children were able to expend creative energies on their posttest performance that resulted in scores equivalent to their pretest TTCT verbal performance. Examination of the effects of the ten creative dependent variables (see Tables 1 and 2) shows a significant univariate effect for verbal flexibility for the OM x Time interaction. Because this was not accompanied by an overall global multivariate effect for OM x Time for the ten dependent variables, this finding can only be mentioned as a trend; that is, of the three verbal TTCT measures, flexibility was most affected by the OM treatment. It is likely, therefore, that the OM-hi children were better able to resist the detrimental motivational effects on TTCT performance of desiring to work on their regular enrichment activities rather than the assigned test battery. The non-OM group appeared to profit (probably from the creativity training in the overall enrichment program activities) to a greater extent than the OM-lo group.

The size of the effect for OM (.07 for Similes, .08 for the mean verbal TTCT) was a smaller effect size than those tested by Rose and Lin (1984) for CPS effects on the TTCT verbal scores. Cohn's (1985) meta-analysis of effectiveness of creativity training found that test scores are less affected by creativity training when the test tasks are dissimilar to the tasks which were used in the training. As described in Chapter I, III, and Appendix C, the CPS training in OM uses some techniques which may be less similar to classroom exercises in divergent thinking than CPS training programs which follow the Osborn-Parnes model.

Affective Hypotheses

Six of the seven hypotheses for the affective construct specified that a significant interaction would occur. The hypotheses were stated in a nondirectional manner for the following predicted effects:

2. There will be a significant interaction on the affective construct scores with respect to effort in OM and time.

4. There will be a significant interaction on the affective construct scores of students in relation to the amount of their experience in the enrichment program and time.

6. There will be a significant interaction on the affective construct

scores of students in relation to the amount of their prior experience in OM and time.

8. There will be a significant interaction on the affective construct scores when these scores are compared with respect to effort in OM, experience in the enrichment program, and change in score from pretest to posttest in the experimental year.

10. There will be a significant interaction on the affective construct scores when these scores are compared with respect to effort in OM, prior experience in OM, and change in score from pretest to posttest in the experimental year.

11. There will be a significant difference in sensitivity of the scores of the univariate dependent variables of the affective construct to student effort in OM.

12. There will be a significant difference in creative self-concept scores with respect to student effort in OM and time.

The mixed model MANOVA did not show a significant interaction for OM and time. There were pre-existing differences in affect among the OM groups, thus supporting hypothesis 11. However, when the posttest scores were adjusted for pretest scores, as well as for three important uncontrolled variables--sex, length of experience in the enrichment program, and amount of prior experience on an OM team--the treatment effect of OM demonstrated significant differences between groups on affect (see Tables 6 and 8, and Figures 7 and 8). The global effect was supported by univariate differences in general self-concept and creative self-concept. There were significantly lower posttest scores in general self-concept (Sears) for the OM-lo effort group than for those students who were not in OM. It is possible that low effort students may have felt their ideas, when offered, were not listened to, and, therefore, may have felt less secure about themselves if they had experienced some of the group to be possibly judgmental. However, there were higher posttest scores in creative self-concept for the OM-hi effort group, thus supporting hypothesis 12. The adjusted posttest scores for Ideas were significantly higher for OM-hi than for OM-lo children (post hoc contrast, $t=-2.57$, $p < .05$). These results suggest that students who are judged by themselves and their peers as low in effort on an OM team suffer a lowering of general self-concept. It is hoped that the lowering of self-concept in the OM-lo effort group was a temporary decrease in self-concept scores attributable to comparisons with peers based on a situational grouping in which students felt the others to be more capable than themselves. Coleman and Fults (1982) noted that some students in gifted programs showed a somewhat lower Piers-Harris general self-concept score during the year of participation in a gifted program. However, the scores increased again 18 months later showing higher self-concept than their less gifted peers. The overall general self-concept was found by Coleman and Fults nevertheless to be above the norm for their grade. The finding in this study of lower self-concept in the low effort group, therefore, suggests that further research of a more rigorous nature is needed to investigate the relationship of OM effort and self-concept. The Sears instrument shows good reliability and appears to be sensitive to treatment effects in other research (Whitmore, 1980) as well as in this study. The effect for creative self-concept suggests that students who exert effort in their work on an OM team may profit from the experience in feeling that they are more creative. The implications of these findings are discussed below.

Main Effects of Time

Creative Construct

A significant effect was found for the change over time on seven of the ten measures of creativity (see Tables 1, 2, 3, and 4). Consistent with findings of other investigators (Rose & Lin, 1984), the greatest changes were noted in verbal and figural originality when time was considered as a main effect. Scores on Similes and five of the six figural measures were higher on posttesting than on pretesting. There was a marked increase in Similes and in figural fluency and originality, but a decrease in elaboration. The univariate ANOVAs of the cognitive creative data showed strongest changes in originality, downward for verbal originality and upward for figural originality. Unlike many other studies, scores on the verbal TTCT declined from pretest to posttest. This finding may indicate a need for caution in interpreting these results. It has been found that performance on tests of creativity are sensitive to environmental factors (Amabile, 1983; Torrance, 1987). When children are removed from favored activities for creativity testing, creativity performance as measured by that test is likely to be depressed (Elkind et al., 1970). This author interprets the significant decline of posttest scores in this study to be such a context effect. It is also likely that ceiling effects of statistical regression toward the mean for posttesting may have occurred in this sample of gifted children where many are highly creative (as evidenced in this study by inordinately high pretest scores). The results show that growth occurred in figural creativity (with the exception of elaboration) and in Similes but not in verbal creativity as measured by the TTCT.

<u>The context effect of the environment on creativity test scores</u>. Similes and the figural TTCT preceded the verbal TTCT in order of administration of the tests in this study. The verbal battery was probably perceived as a challenging and appropriate activity for enrichment students during the pretesting, but during the posttesting appeared to be perceived as repetitive and as interfering with student time needed to complete preferred enrichment activities for end of the year credit. Moreover, by the end of the year, new students had been added to most of the enrichment classes. These new students had not been involved in the testing for the study and were therefore free to work in enrichment class learning centers

on other activities while the children in the study were tested. This may have added to the negative attitude of those being tested.

Elkind et al. (1970) clearly demonstrated the context effect, that students perform significantly better on the Wallach-Kogan creativity tests when the ongoing activity which was interrupted was "uninteresting" to them than when the ongoing activity which the testing session interrupted was one that was intrinsically interesting to these same children (see discussion in Appendix E). Because of the strong evidence from prior research that increases in TTCT scores occur after training for creativity and because of creativity training in the enrichment program and teacher observations of consistent increase in children's creative behavior in the program, an increase in posttest TTCT scores was expected. However, because the tests were given in a setting in which favored activities were perceived to be significantly interrupted for lengthy testing, and because scores increased on Similes, a 15-minute test which was administered as the first test in the lengthy battery, the decrease in scores on all three verbal TTCT measures and on figural elaboration is therefore considered to be primarily attributable to a potent context effect. Because the reduction in creativity scores was most noted for third grade (see Tables P-9 and P-10 and Figure P-11, Appendix P) it is possible that a history effect may have contaminated these results. That is, extensive Metropolitan Achievement Testing had only recently been completed by the third graders (see limitations, p. 125), and thus may well have contributed to lowered motivation on the TTCT verbal posttests.

As seen by the review of the literature on creativity, we are dealing with a very complex, multi-faceted phenomenon. The overwhelming consensus is that training increases creative thinking performance on the TTCT, particularly when the training utilizes the Osborn-Parnes procedures of CPS (Feldhusen & Clinkenbeard, 1986; Rose & Lin, 1984; Torrance, 1972a). However, such was not even clearly the case in this study. The training in diverse aspects of creativity, in which all of the children in the study participated during their enrichment class time would be expected, based on prior research, to result in significant gains in posttest creative thinking, and especially so in verbal originality and the other verbal measures (Rose & Lin, 1984). The reverse, however, occurred; verbal scores decreased significantly for the children in this study and decreased most significantly for verbal originality.

It is likely that children who are creative could be particularly susceptible to the motivational impact of environmental effects. Scores of highly creative children on the TTCT are more responsive to the effects of cue-rich versus cue-poor testing environments than are the scores of children who are low in creativity (Mohan, 1970, cited in Torrance, 1972a). Many of the children in this study were creative as well as gifted, as may be seen by the high pretest scores on the creative measures (see Tables 1 and P-2, Appendix P, and Chapter IV, Results, pp. 141-145). Therefore, the significant decline of posttest scores is interpreted as reflecting the motivational impact of the environmental context. Some reduction of extremely high pretest scores was probably further affected by the phenomena of regression toward the mean (see Chapter IV and discussion below).

Affective Construct

A significant multivariate effect of change over time was found in the construct of affect across all subjects in the study. This overall change was mostly attributable to an increase in creative self-concept (Fishkin, 1987a, 1987b) and locus of control for negative events.

The creative self-concept scale showed a significant increase over time across groups (see Tables 6 and 7, and Figure 6), suggesting that it measured aspects of creativity fostered within the gifted program. However, the size of this effect is very weak, N^2=.02. The newly developed scale demonstrated an acceptable concurrent validity correlation with the general self-concept scale because creative self-concept was more highly correlated with overall self-concept than with the locus of control measures. The increase in creative self-concept, but not in general self-concept, over time is consistent with the literature on specificity in regard to academic self-concept (Shavelson & Bolus, 1982). The increase in locus of control for negative events across groups suggests that an effect of the overall gifted program is to support responsibility for one's own actions as a self-directed learner. The occurrence of change in the specific areas of affect of creative self-concept and of attribution for responsibility is consistent with the curriculum goals of the gifted program per se which stresses creative thinking and self-directed learning.

The failure to find the hypothesized interaction between time and effort in OM in the repeated measures analysis may be due in part to the rather potent treatment effect experienced by all of the subjects, including the control group. However, in the discussion above for the major hypothesis, when some of the variance attributable to uncontrolled variables in this quasi-experimental study was accounted for by means of MANCOVA procedures, the effect of the OM treatment condition was indeed apparent.

Significant changes for time across groups occurred in creative self-concept and locus of control, but not in general self-concept. These changes, when tested by strength of association, were found to be of a very small effect size. The results indicate that growth occurred in creative self-concept, internal locus of control, and creativity, with the exception of performance on the verbal TTCT, across all students in the program.

Main Effects for Pre-Existing Differences Between Groups

Creative Construct

Pretest differences in Odyssey of the Mind groups. The strong pre-existing differences between the OM groups is shown by the significant differences for Similes, closure, and the total of creative strengths (see Tables 1 and 2, and Figure 2). The pretest means for the non-OM group are lower on these measures than those for students who chose to be on an OM team. Moreover, on Similes and creative strengths, scores were lowest for the non-OM group and highest for the OM-hi group. These results demonstrate that the groups were not equivalent in scores on creativity measures prior to experiencing the experimental treatment, effort on an OM team, during the year of study.

The cluster of creativity variables which significantly differentiates those children who decided to join an OM team from those who did not should be noted. The variables which were much higher for the children who chose to be on a team were Similes, closure, and the creative strengths. This pattern of creativity variables is consistent with those variables which reflect the two experience variables in the study. The patterns of these measures to each other and to the treatment effects of this study are discussed more thoroughly below under relationships of the different dependent variables to the treatment effects and under criterion validity in Appendix E.

Pretest differences for experience in the enrichment program and in prior experience on an Odyssey of the Mind team. As noted in Chapter IV, the pattern for these two effects was remarkably similar. There was a significant main effect for both experience variables on the creative construct for the analyses of the full ten variables, and for Similes and the figural Creativity Index for the analyses of the abbreviated creative construct.

The pre-existing differences between subjects in this sample, when they were divided according to either of the experience variables, experience in the program or prior experience on a team, showed this pattern in the dependent variables: students who were more experienced were lower on verbal flexibility, $p < .05$, but higher on total of creative strengths, $p < .001$, closure, $p < .01$, Similes $p < .01$, figural originality, $p < .05$, and figural elaboration, $p < .05$ (see Tables P-4, P-5, P-6, and P-7 in Appendix P). Some possible interpretations of the relationship of the different creative measures to treatment effects are offered below (see pp. 198, 220-221).

Affective Construct

The relevant independent variables of experience and pretest differences between OM groups were distributed more equally within this sample in their effects upon the affective construct than upon the creative construct. Of the three independent variables under investigation, only the OM effort group variable showed significant pretest differences; the two experience variables did not. However, unlike on the creative data, the experience variables of months of experience in the enrichment program and of years of prior experience on an OM team were important predictors of posttest affective scores, as shown by the regression equations (see Table P-23, Appendix P) and were, therefore, included as covariates in the MANCOVA procedure (see Table 8). Although not supported by an associated global multivariate effect, two possible differences were noted, however, on these experience variables. The stepdown analyses indicated that there might be a significant difference in locus of control for negative events between groups who were experienced in the program and those with less than a full year of experience I-, $F(1,103)=4.00$, $p < .05$) (see Tables P-19 and P-20, Appendix P). A comparison of the mean scores shows the students who were experienced in the program had a greater sense of responsibility for negative events ($\overline{X}_{I-}$ = 10.32) than students new to the program ($\overline{X}_{I-}$ = 9.19. For the variable of prior experience on an OM team, the stepdown analyses found higher creative self-concept in students who were experienced on an OM team ($\overline{X}_{Ideas}$ = 34.70) than in those who had no prior experience ($\overline{X}_{Ideas}$ = 33.98, $F(1,102)=4.37$, $p < .05$) (see Tables P-21 and P-22, Appendix P). Because these findings are not supported by a significant global effect, it is possible that they are chance effects. It is, however, possible that the power of the tests was reduced by the low sample sizes in the cells: less than 13 months experience group ($\underline{n}$=17, 9, and 11) and ($\underline{n}$=10, 11, 12, or 37) of the cells in the Prior Experience in OM x OM analysis. It is likely that such reduced power can produce "a nonsignificant multivariate F, but one or more significant univariate Fs" (Tabachnick & Fidell, 1983, p. 231). Therefore, future researchers should expect to need a larger sample of subjects.

There was a significant low level global effect of OM on the pretest affective construct. This low level effect, although not supported by the univariate analysis, was supported by a significant stepdown F statistic for creative self-concept (see Tables 6 and 7). These results showed that students choosing to join an OM team held a higher view of themselves as creative persons than the children who chose not to join OM. Although the univariate strength of association estimate of this effect showed it to be at a low level, N^2=.03, the results corroborate observed differences between children who elect to join a team for creative competition and those who do not. Because these findings are congruent with the main effect of time, an increase of scores in creative self-concept and in I-, these stepdown findings are interpreted as evidence in support of the effect of the enrichment program to increase qualities consistent with self-directed learning goals.

Summary of Findings Bearing on Relationships Among the Different Dependent Variables

In this section, the important effects on the different variables will be briefly discussed, most of these effects are more fully discussed in other sections of this chapter. The effects listed in this section are those that, by the N^2 strength of association measure, were found to be important. That is "judging from the present state of the art in the behavioral sciences, any time {one} can account for more than 10% of the variance, {one is} doing better than the vast majority of studies" (Linton & Gallo, 1975, p. 331).

Creative Construct

The multivariate global effects noted for the creative construct--the interaction of OM and time, the pretest differences between OM groups, the pretest differences between children in the study who were experienced in the program or who had prior experience on an OM team and those who had less experience, and the overall effect of time, as well as differences between grade levels--all showed strong relationships for some of the variables in the construct. The important univariate results of these effects are presented below.

The predicted effect of an interaction of OM and time was significant with the multivariate strength of association at N^2=.11. With control through covariance of the effects attributable to the uncontrolled variables of grade and sex, the effects in the MANCOVA analysis of adjusted posttest scores, which were accounted for by the difference between the OM groups, rose to .07 for Similes and .08 for mean verbal TTCT scores. These results, therefore, indicate that the obtained results of OM CPS training were a significant but low level effect.

The univariate effects for creativity which were greater than N^2=.10 occurred most often for verbal and figural originality, Similes, and the total of creative strengths. The proportion of the total variance on Similes that was attributable to pretest differences between OM groups was N^2=.11. The proportion of total variance in the analysis examining the effects of prior experience on an OM team on the pretest scores accounted for .12 of the total variance in Similes and for .14 of the total variance in the total of creative strengths variable. Among the strongest univariate effects in this study was the overall effect of time in the mixed model 3x2 MANOVA for the effects of OM and time; differences between pre- and posttests accounted for .20 of the variance as an increase in figural originality, and .21 of the variance as a decrease in verbal originality. The decrease in scores for verbal originality is discussed above (p. 196) and in Appendix E as in part likely to be attributable to detrimental motivation of a context effect in this study.

Very large effects were noted for the effect of grade, which had not been planned as a manipulated or controlled variable in the study. When it became apparent that there were large differences on several of the dependent variables attributable to differences in grade level, it was necessary to test the strength of the effects of grade by using it as a blocking variable. The interaction of grade and time accounted for 18% of the total variance of closure. As noted below (p. 225), the verbal scores reflected very significantly lower scores for Grade 4 than for Grade 2+3. Although significantly lower scores of Grade 4 may be attributable to the "fourth grade slump" in creative thinking (Torrance, 1974b; Williams, 1976), some of the very large differences at Grade 3 and 4 in this study and in Torrance and Safter (1986) could be explained by the difference between motivational effects for group versus individual administration of the verbal tests for Grade 3 and younger.

Although closure consistently reached significant levels of findings for the effects of each of the pre-existing variables, it resulted in a strong association only with effects for grade and grade by time. Unlike Similes, closure did not show the strength of association above .10 with the effects of differences attributable to prior experience on a team, previous experience in the enrichment program, and differences between children who did or did not join an OM team.

As discussed above (p. 199) and in Appendix E, closure, creative strengths, and Similes appear to be the creativity variables associated with experience in the enrichment program, as well as with prior experience on an OM team. The ability to resist premature closure is considered by Torrance (1979) to be an important ability for productive creativity because it is associated with psychological openness. It is possible that the difference in scores on closure could reflect the training in learning to defer judgment that is taught the children in the enrichment program and on OM teams. This interpretation is borne out by the regression equation (Table P-8, Appendix P) in which years of prior experience on an OM team was a significant predictor of the equation for closure.

Because closure and the creative strengths were higher as pre-existing differences between OM groups and experience groups, it is surprising that they did not show a greater increase in the OM-hi group over time in comparison to the other OM groups. Actually, inspection of cell means shows that was indeed the case; of the figural variables, closure and creative strengths showed the pattern of increase for the OM-hi effort group and almost no growth for OM-lo and non-OM groups (see Table 2). These effects, however, did not reach a significant level.

Affective Construct

In general, the affective effects in this study accounted for lesser amounts of the variance in the respective analyses than the effects for the creative construct. The multivariate effects for pre-existing difference between OM groups and the effect of time across groups accounted for a small but important amount, .14 and .16, of the variance in the affective construct for the mixed model analysis. For the MANCOVA, the effect of OM on adjusted posttest scores accounted for .17 of the variance of the multivariate construct for affect. Of the univariate effects in the construct, change from pre- to posttest on creative self-concept across groups accounted for only .02 of the variance in the linear combination of scores for affect. The differences between OM groups in general self-concept accounted for .09 of the variance and in creative self-concept accounted for .07 of the variance in the MANCOVA test of the adjusted posttest scores. However, none of the studies which were reviewed that investigated change in affect associated with creativity training found any significant changes in self-concept (Fults, 1981; Kolloff & Feldhusen, 1984; Tadlock, 1981). It would, therefore, seem from other research that it has been difficult to measure the relevant changes in affect, if indeed such changes occur in relation to creativity training. Therefore, given the state of the art in obtaining changes in affect that reach significance but account for less than 10% of the variance, it is possible that the changes in affect may simply be an accurate reflection of an effect of very low power. (For further discussion of the findings in relation to these affective measures see below, pp. 221-228, 243-245.) These findings support hypothesis 12, showing a significant difference in creative self-concept in relation to effort in OM and time.

Findings Relevant to Methodology of Creativity Research

Comparisons of Statistical Procedures

Two methods of adjusting the dependent variables for pre-existing differences in creativity and of affect were used. The first, a mixed model analysis, permitted assessment of changes in individuals over time in relation to the OM effort independent variable. For this analysis, the predictable variance removed from the within subjects error term accounts for the contributions associated with each individual subject's unique variability on the various dependent measures. Such repeated measures analyses, however, are a low power, conservative technique which can inflate the Type II error rate (Tabachnick & Fidell, 1983). The possibility of increasing the power to test for differences between groups which may have been suppressed in the mixed model analysis contributed to the decision to use MANCOVA as the second method of analysis to explore the nature of the findings.

Because there were several variables impinging on the sample in this study which could not be controlled by random assignment of subjects to the independent groups, analysis of covariance was used as a statistical matching procedure. Interpretation of results of MANCOVA based on means adjusted as if all subjects in the group had the same pre-existing scores on the covariate, however, should be made with some caution. Results obtained from ANCOVA should not be used to infer causality of the treatment condition to obtained differences between groups (Tabachnick & Fidell, 1983, p. 180).

In assessing the initial analysis of covariance, a violation was found of the assumption of homogeneity of regression. That is, the slope of the regression lines of some of the dependent variables on some of the covariate measures was not equal across cells of the independent variables. For example, in testing this assumption for the affective MANCOVA, an interaction was found between OM and pretest locus of control measures, grade, and years of experience in OM. Grade was not an important predictor in the affective regression equation (see Table P-23, Appendix P) and was, therefore, eliminated from the subsequent MANCOVA procedure. However, MANCOVA is not robust to violations of homogeneous relationship between the covariates and the dependent variables in the different groups of the analysis. When adjustment of the dependent variables is performed by covariates which interact with the different groups of the study, then "MANCOVA is an inappropriate analytic strategy, both statistically and logically" (Tabachnick & Fidell, 1983, p. 234). Because the essential covariates of pretest I- scores and years of prior experience in OM yielded significant interaction with differences between OM groups, the results of the affective MANCOVA must be viewed with caution.

The analysis of covariance model was used to examine the effects of the OM independent variable after removing the effects of the covariates on the posttest dependent variables. This analysis clarified the relationship of OM to the creative and affective dependent variables. The MANCOVA procedures for the creative construct satisfied the necessary asumptions for covariance analysis. By removing the uncontrolled variance accounted for by the covariates of grade and sex, the MANCOVA confirmed the significant finding of the effect of OM on adjusted posttest mean verbal TTCT scores of the abbreviated creative construct and revealed a significant effect on Similes as well. The MANCOVA of affect revealed a significant positive effect on creative self-concept for the high effort OM group. However, a detrimental effect appeared on the adjusted posttest scores for the Sears general self-concept variable.

Two statistical transformations were performed on the creative data: the ceiling was set at three standard deviations above the norm of 100, and the nine dependent variables of the figural and verbal TTCT batteries were combined to create only two dependent variables plus Similes. The effects of setting the ceiling of TTCT standard scores at 160 was to control the undue effects of extreme mean verbal TTCT scores in 17 of the cases as discussed in the section below.

The data for the ten dependent variables were analyzed in order to obtain information about these new variables (see discussion in Appendix E on contributions of this study to validity information on the TTCT). However, the verbal TTCT measures showed excessively high intercorrelations, r=.76-.95 (see Table P-1, Appendix P). The magnitude of the within-group intercorrelations can affect the power of the MANOVA test (Bray & Maxwell, 1985). In repeated measures analysis, the number of dependent variables for each outcome variable is multiplied by the number of times the repeated variable is measured for each subject. Therefore, where there are several dependent variables in it, the repeated measures analysis can reflect a greater loss of power than a simple factorial design (Harris, 1975). Because of the high intercorrelations, the unequal $\underline{n}$ and mixed model design, and a determinant of the covariance matrix indicating an extremely low value, there were strong statistical benefits toward deleting variables or using some kind of combination of scores. Rather than deleting any variables or using principal components analysis (Tabachnick & Fidell, 1983), the verbal scores were averaged and the figural Creativity Index recommended by Torrance & Ball (1984) was used. The transformation of measures from 10 dependent variables to three yielded some improvement of the determinant and permitted the patterns in the treatment effects of the study to emerge more clearly (see Tables 1 and 4).

Instruments

Creative Measures

Streamlined Scoring and the Torrance Tests of Creative Thinking

Effects of cut-off (ceiling) transformation of standard scores. Some of the scores were transformed downward by setting the ceiling for any TTCT standard scores at less than or equal to 160, the equivalent of three standard deviations above the mean of the standardization group. Multivariate analysis procedures are sensitive to the effects of outliers and therefore procedures such as transformation or deletion of cases are recommended to reduce the effects of extreme scores on the analysis (Tabachnick & Fidell, 1983). In examining the 10 separate creative dependent variables, 47 students had one or more of their eight TTCT standard scores at greater than 160. However, the mean verbal scores of only 17 students were affected by the ceiling transformation and no Creativity Index scores reached the level of 160.

A comparison of the cell means of the creative construct before and after transformation (see Tables 2 and P-2) showed that the most pronounced changes were in all cells of the verbal pretests and the pre- and posttest cells of figural originality. The mean scores for figural originality were lowered for most groups with a marked reduction of posttest scores for seven students who were OM-lo with less than 13 months experience in enrichment, for seven OM-hi with more than 13 months experience; and for six students who were OM-lo with no prior experience on an OM team, and for six OM-hi with prior experience on a team. The standard deviation of the cells which contained transformed scores were therefore more equivalent to the standard deviations throughout the data.

As noted in Chapter IV, there was a significant decline in verbal TTCT scores from pre- to posttest in this study. As discussed above (pp. 194-196) and in Appendix E, the decline in verbal TTCT scores may be, in part, attributable to lowered motivation related to a context effect (Elkind et al., 1970) of the children viewing the TTCT test activity as an interruption and an interference with their preferred activites in the enrichment program. Moreover, of the 17 children whose mean verbal scores were adjusted, 10 children whose pretest scores were inordinately high showed a very noticeable decline on their posttesting scores which may have reflected regression toward the mean. Statistical regression toward the mean upon posttesting becomes more pronounced when the groups are selected on a given variable which correlates with the dependent variable (such as intelligence and creativity) and when the initial scores are more extreme (Isaac & Michael, 1981). Moreover, the more highly creative children were probably more likely to be affected by environmental considerations, such as testing during enrichment class time, than those whose scores were closer to the average (Mohan, 1970, cited in Torrance, 1972a).

A comparison of the results of the MANOVAs before and after the scores were transformed by setting the ceiling at less than or equal to 160 shows only minor changes in the effects of the analysis. The main effect of the transformation was to reduce the univariate effect of OM x Time on verbal flexibility. Although the transformation reduced the effect on verbal flexibility of the major hypothesis for this study, the OM x Time interaction was more pronounced on the analysis with only three creative dependent variables than with 10 variables. Examination of the statistics for the within-cell correlations of between-subjects effects (Bartlett's test of sphericity) (SPSS, 1986) and within-subjects effects (Mauchly's test of sphericity) (SPSS, 1986) showed that the analyses of 10 dependent variables were the most suspect of violating the assumption of homogeneity of the variance-covariance matrix. The two transformations produced a set of dependent variables which enabled the results of these complex analyses to be interpreted with appropriate confidence (see Appendix B) (Bray & Maxwell, 1985; Harris, 1975; Tabachnick & Fidell, 1983).

Raw scores versus standard scores analysis. Normally, when standard scores are available, the standard scores are preferred to percentile scores or even raw scores as the most powerful scores for statistical use because they are generated from the properties of the normal probability curve (Isaac & Michael, 1981). The use of the standard scores was recommended by the latest manual for streamlined scoring of the TTCT (Torrance & Ball, 1984). In the manual for the standard scoring (1974b), Torrance recommended that, if any kind of composite or total score was to be used, the standard score conversion was the appropriate measure. In this study, the conversion to standard scores was beneficial for two purposes: it enabled a comparison of the relative strengths of the various TTCT measures for the students in the sample, and it enabled computation of the composite scores for mean verbal TTCT performance and for the figural Creativity Index. As stated in the preceding section, the use of these two composite scores provided a more stable basis for analysis of the data. Moreover, the use of the composite scores provides a more reliable measure than does use of scores from the separate variables (see discussion in Appendix E and Torrance, 1974b).

There appear to be disadvantages to standard scores, however, when used to test the effects of an experimental treatment. The conversion to the standard scores introduces a source of unreliability, particularly to the scales for closure and for titles because of the limited range of these subscales (O. F. Anderhalter, personal communication, June 15, 1988, see Appendix A and Appendix E). Therefore, the raw scores, rather than the standard scores, are recommended for estimates of reliability for the TTCT. The improvement in TTCT inter-rater reliability estimates for raw scores over standard scores is graphically shown in a comparison of Tables F-6 and F-7 in Appendix F. A further disadvantage for use of the converted standard scores lies in the possibility that the sample under study may frequently consist of a disproportionately high number of creative individuals. The scores for the sample in the present study were markedly skewed toward high creativity (see Chapter IV, pp. 141-145). True standard scores are an inappropriate measure when the data is markedly skewed (Isaac & Michael, 1981). The effects of transformation of the raw scores to the grade-level based standard conversion scores are discussed more fully in Appendix E. This author's recommendations, based on the findings of this study, are that the use of composite scores for the TTCT provides a more appropriate measure than use of the individual subscale scores. This is true for the verbal measures because of the duplicative amount of correlation shared by the three verbal measures (Chase, 1985). The composite scores of the Creativity Index are preferred for the streamlined battery because of the increased reliability of the composite score and the use of the highly appropriate creative strengths. The composite scores should be more useful for purposes of identification than the subscale scores alone because of the questionable reliability of the closure and titles scales when converted to standard scores (see Appendix E and Tables F-6, F-7, F-8, and F-9, Appendix F). However, when the purpose is to examine different qualities of creative thinking assessed by the different measures, then standard scores should be the preferred measure. In highly skewed samples (see Table F-10, Appendix F), the raw scores probably ought to be used as the preferred unit for analysis.

Reliability. In the process of ascertaining the reliability for the streamlined figural scoring for this study, a study within the study emerged. The obtained inter-rater reliability is reported in Chapter III (pp. 111-113) and Appendix E. The results of these comparisons yielded the following inter-rater reliability coefficients for the sample in this study. The 119 verbal A booklets scored by two professional scorers of Scholastic Testing Service obtained coefficients for fluency=.95, flexibility=.86, and originality=.86. There were 20 figural B booklets randomly selected for the inter-rater comparison with subsequent correlations by this scorer and by the senior professional scorer of the Scholastic Testing Service: fluency=.99, originality=.96, titles=.82, elaboration=.76, closure=.89, and total of creative strengths=.66 (see Table F-8, Appendix F). These reliability coefficients are very respectable but are considerably lower than those reported in the manual (Torrance & Ball, 1984). The results of this TTCT reliability study within the larger study offer evidence to confirm the conclusions that it is important to use a single rater and to develop local norms based upon the scores from that rater (Halpin & Halpin, 1974). These findings confirm those by Halpin and Halpin that the greatest discrepancy occurred on elaboration. For purposes of comparison of individuals for identification for a gifted/talented program, it would especially be important to use a single rater (Rosenthal, DeMers, Stillwell, Graybeal & Zins, 1983; J. D. Kauffman, personal communication, November 11, 1987, see Appendix A).

Alternate forms reliability coefficients were computed for all of the measures in this study, including the TTCT (see Table F-9, Appendix F). These coefficients ranged from .18 for closure to .51 for verbal flexibility, and .69 for the Sears. As noted in the discussion in Appendix E, the test-retest reliability correlations obtained for this study are considerably below the values of alternate forms reliability previously reported (Torrance, 1974b; Torrance & Ball, 1984). However, no previous alternate forms reliability comparisons were reported for closure, for creative strengths, or for the Creativity Index. The reliability coefficients obtained in this study should be viewed as underestimates for several reasons. The estimates were based on standard scores for the three verbal measures and the five norm-referenced figural measures, rather than on raw scores; and as stated above, and in Appendix E, TTCT reliability is inappropriately lower when computed on standard scores than when computed on raw scores. In addition, the retest reliability measures were separated by a five-month interval with intervening significant treatment effects. Furthermore, the reliability study was conducted in a sample of gifted students, and the resulting scores were likely to be limited in distribution to the upper ranges of creativity. Moreover, regression toward the mean appears to have occurred because many of the extremely high pretest scores showed disproportionately large decreases on the three verbal measures and possibly on elaboration as well.

Validity. The results of this study add to the presently still sparse base of validity information available on the relatively new streamlined scoring for the figural tests. The previously available validity information on the streamlined procedures is briefly discussed in Chapters II and III and in Appendix E. The additional evidence of validity is presented in Appendix E and, therefore, is only briefly presented below.

Examination of the intercorrelations of the TTCT variables (see Table P-1, Appendix P) revealed that, with the new streamlined procedures for scoring elaboration and new measures of abstractness of titles and resistance to premature closure, the figural measures each appear to offer unique qualities to the construct. The high level of intercorrelation with fluency that is still present in the verbal battery (Chase, 1985) is no longer a problem with the figural battery. Originality is most closely correlated with fluency, reflecting the TTCT definition of originality as ideas that are statistically infrequent or unusual. However, originality also is significantly correlated with closure and with the creative strengths. The creative strengths bear significant correlations with each of the measures of the figural battery and with Similes. The streamlined scoring adds a richness or essence of qualities of creativity to the scoring dimensions that are lacking in the divergent thinking measures of fluency, flexibility, and elaboration, as is evident from the lack of correlation to the creative strengths with any of the verbal TTCT measures. Criterion validity is shown by the relationships of the creative strengths, elaboration, closure, and verbal originality to Similes. This author considers that the scores obtained from Similes reflect exceptional quality and capture the richness of verbal creativity to a greater extent than do the divergent thinking scores obtained from the verbal battery.

The findings from the multiple regression equations, summarized in Tables P-8, P-17, and P-27 (see Appendix P), support the relationships noted in the correlational data. The new measures of abstractness of titles, closure, and creative strengths, and the resultant Creativity Index bear greater relationship with each other and with Similes than to the verbal TTCT measures. The Creativity Index carries a significant prediction for figural fluency, $\beta=.25$, but the score is far less dependent on fluency than is the mean verbal measure, $\beta=.53$ (see Table P-17). The regression equation results, then, add to the understanding that the streamlined measures improve upon the divergent production measures by reflecting a greater breadth of creative behaviors.

The responsiveness of the figural TTCT measures to the treatment effects of this study shows primarily in the pre-treatment effects noted on the pretest scores. There was a cluster of creativity measures which discriminated between the children who had the greater experience with creativity training than those who did not. These effects were noted on two experience variables, experience in the enrichment program and prior experience on an OM team. Pre-existing differences between children who sought additional creativity experiences by their willingness to join an OM team also clustered with some of these same variables. The cluster of indicators which seemed most strongly associated with the greater amounts of experience with creativity training and with the propensity to join an OM team consisted of higher scores on closure, Similes, and total of creative strengths. In addition, the children who had greater experience in enrichment and/or on an OM team scored higher on figural elaboration and figural originality, but, surprisingly, lower on verbal flexibility than did those with less experience. Strength of association measures indicated that the strongest of these effects were Similes and the creative strengths.

Resistance to premature closure is considered by Torrance (1979) as the creative ability which taps psychological openness. He recommended that the prime technique for training this creative ability is learning to defer judgment, as taught in the CPS process model (Parnes, 1972). Deferring of judgment during brainstorming and during some affective training activities as well is a principle stressed in the enrichment classes and taught to the coaches during their training session (see Appendices C and D). Years of prior experience on an OM team was a significant predictor of the regression equation for closure (Table P-8, Appendix P) thus offering further evidence that the higher scores on closure associated with greater experience in CPS training suggests that CPS training was beneficial to increasing creativity.

Figural originality and verbal originality were most affected by the effects of time; there was a strong increase in figural originality across all groups in the study which probably reflected the effects of creativity training in the overall enrichment program. These results are consistent with the meta analyses by Rose and Lin (1984) and a review of creativity research by Cohn (1985). Rose and Lin suggested that increases in verbal creativity were associated with CPS training and increases in figural creativity were more likely to be associated with training by diversified enrichment programs. Cohn's results indicated that creativity training can increase the fluency and originality of creativity test scores. However, Cohn also found that differences in warmup prior to the test and/or in other motivational factors can affect the scores to the same degree as extensive creativity training. Creativity test scores were more likely to be increased on those test tasks which are like the training tasks than test tasks which are dissimilar to the training. The figural tasks of the TTCT were similar to some of the training activities in the enrichment program, such as activities from New

Directions in Creativity (Renzulli, 1973), but dissimilar to most of the OM creativity training.

As noted above under main effects of time and in Appendix E, the verbal scores showed a significant decline from pre- to posttest. These results may likely be attributable to the presence of the strong context effect which seriously depressed the verbal posttests. The deleterious motivation noted during the verbal posttesting was not readily apparent on the posttesting for Similes or the figural TTCT, which preceded administration of the verbal posttests. However, the figural elaboration scores, unlike the other figural measures, showed lower scores on the posttesting. Torrance (1979) suggested that willingness to elaborate is an important creative ability which taps the person's willingness to expend effort and develop creative ideas. The lowered scores of elaboration on the figural posttests was probably also due to the context effect, lowered motivation to expend creative energy on the test, which although still somewhat novel, was no longer as intrinsically interesting to many of the children as completion of some desired enrichment class activities for end-of-the-year credit.

As in the present study, Cohen (1987) found that the students in OM were higher in creativity on the verbal TTCT than on the figural. She found that the two teams which became championship teams at the level of world final competition were higher than the other three teams on every subscale of the verbal TTCT and had a majority of their team members who had verbal scores greater than the 90th percentile. Both of the winning teams had their highest figural scales as fluency and originality. Cohen's study, however, only used descriptive statistics so no meaningful statements about extent of comparisons among the five teams can be made.

Value of the Creativity Index. As discussed in Appendix E and in the preceding section on reliability, the Creativity Index was the most reliable, along with figural originality, of the figural measures in this study as assessed by alternate forms reliability (see Table F-9, Appendix F). The Creativity Index provides a measure to utilize the creative strengths in a meaningful composite score. The creative strengths, as discussed above, appear to show good criterion validity with Similes (Tables P-1, P-8, and P-27, Appendix P) and good validity in the cluster of creative abilities which helped to distinguish between children who had prior experience in creativity training and those who did not (see Tables P-4, P-5, P-6, and P-7, Appendix P). Despite the significant decrease in elaboration across groups, the Creativity Index showed a significant increase across subjects from pre- to posttesting, which may be attributable in part to maturation and/or may reflect the creativity training effects of the enrichment program (see Tables 3 and 4). These results were consistent with those of Gray (1986) who found that the Creativity Index increased with subscale increases on fluency and originality but with decreases in elaboration. Gray, however, also found decreases in closure and abstractness of titles in the children in her study who received training in creative drama. The decrease in closure in children who received training in dramatics in contrast to the increase in closure for the children on OM teams in this study supports the validity of closure as a measure which is possibly associated with training in CPS.

Implications for the Verbal Measures

As noted above and in Appendix E, there was a significant decline in verbal scores on the posttests. This author considers that there was ample indication of a potent detrimental context effect in that the testing was interrupting a special class that was highly motivating to most of the children. The only group whose verbal scores did not decline were the children who were in the OM-hi effort group. Perhaps these children, who had displayed high effort on their teams, had learned to apply themselves better or perhaps to show a willingness to dig a little deeper into hard work and, therefore, were less adversely affected by the lengthy testing during a favored class activity than were their classmates.

Some comments are pertinent to the use of an abbreviated verbal battery. Torrance (1979) indicated that the use of only one or two of the verbal tasks of the full verbal TTCT battery might reduce the predictive validity of the measures. Some of the abbreviated battery reliability studies obtained as high reliability on the shortened versions as the full battery studies (Torrance, 1974b). However, if an abbreviated battery is necessary, Torrance recommended this priority for inclusion of the seven testing activities: "(1st) Product Improvement, (2nd) Just Suppose, (3rd, 4th, and 5th) Ask and Guess, (6th) Unusual Uses, and (7th) Unusual Questions" (1974a, p. 9). As recommended for the eventual streamlined verbal procedures (Torrance & Ball, in press), activity 6, Unusual Questions, will be omitted from future editions of the verbal battery, and activity 6 was, therefore, omitted from the battery used in this study. In retrospect, this study would have benefitted from a shorter testing battery because the class environment is one in which the children normally have many choices and is normally free of <u>any</u> standardized testing.

The intercorrelations of the verbal scores were inordinately high in this sample, .76-.84 for verbal B, .91 to .95 for verbal A (see Table P-1, Appendix P) when compared to .74 to .80 for seven TTCT measures at sixth grade and .69 to .81 for six TTCT measures at the college level (Torrance, 1974b). Two factors present in this study may in large part be contributing to these higher levels of relationship among the verbal measures. It is possible that the high correlation may be reflecting the nature of the sample in this study; that is, the students in this sample were mentally gifted, with IQs ranging from 130 to 162+, the ceiling on the Stanford-Binet. The scores on the verbal tests were positively skewed with a range from 84 to 231. Seventeen students obtained mean scores on the verbal form A tests that were greater than 160. As reported above, these scores were set at a ceiling of 160.

Perhaps it is possible that the high IQ in this sample of gifted children could contribute to the high intercorrelations of the verbal measures. On the other hand, it is also possible that if there is a relationship between IQ and creativity, it is possible that the restriction in range could have lowered the correlation.

The other factor which may have contributed to the excessively high intercorrelations is that the correlations obtained from this study are based on a set of measures different from the full TTCT scored by standard procedures cited by Torrance (1974b). The streamlined manual for the figural tests provided no tables of intercorrelation of the new streamlined measures with each other and with the verbal TTCT measures (Torrance & Ball, 1984). A comparison of the values in the intercorrelations of the verbal TTCT, the figural scored by streamlined procedures, and also Similes (see Table P-1, Appendix P) with those cited by Torrance (1974b) shows that the figural tests scored by streamlined procedures, with the exception of only figural originality and the total of creative strengths, carry lower intercorrelations with each other than do the divergent production scores produced by the standard procedures. The streamlined scoring procedures (Torrance & Ball) have created a method of scoring the TTCT which contains subscale measures that express more diversity and richness present in the creative behaviors than possible in the outdated method of standard scoring (see Appendix E). Moreover, the intercorrelations and regression equations in Tables P-1 and P-8 (Appendix P) show that the new streamlined measures are more different from each other and from the verbal divergent production scores than are the figural TTCT scores produced from the older scoring method (Torrance, 1974b). The comparatively lower correlations of the six figural scales and the presence of Similes in the matrix, then, explain some of the high intercorrelation in the verbal measure. The main reason, however, is that the divergent production scores are highly interdependent; the originality and elaboration scores are far more dependent on the number of responses produced--that is, the fluency score--than is true for the abstractness of titles, closure, and total of creative strengths measures.

A different intercorrelation matrix was created to test the relationship of the affective measures to the creative measures of this study. In Table P-25 (Appendix P) the intercorrelation matrix of the 10 creative and four affective posttest dependent measures with the respective pretest and posttest variables reveals that the values of the intercorrelations of the verbal posttest creative dependent variables do indeed vary with changes in constituency of the matrix. By adding the four affective variables to the matrix the intercorrelations of the verbal measures were reduced from the values shown in Table P-1 to .83 from .88; these values, however, are still higher than the values cited by Torrance (1974b).

It was surprising that Similes, a measure of verbal creativity, carried stronger correlations with three of the figural measures, elaboration, closure, and the creative strengths, than did any of the verbal measures, including verbal originality (Table P-1, Appendix P). The regression equations (Tables P-8 and P-27, Appendix P) showed that the creative strengths is a significant predictor of Similes. Figural fluency predicted verbal fluency, but was a suppressor variable for flexibiity along with Similes for verbal originality (see Table P-27). The creative strengths and figural originality, however, showed in the prediction equation of verbal originality when the 20 creative dependent variables were considered as predictor variables (Table P-27) without the presence of the control variables (Table P-8). Verbal and figural originality on the TTCT are primarily defined as those responses that are statistically infrequent (Torrance, 1974a; Torrance & Ball, 1984). Although high scores in Similes reflect fluency of, as well as rarity of, the responses, a high score depends on exceptional quality of the response. Responses which have high quality and attributes of evoking emotional responses in the reader are recognized by the Similes scoring and by the streamlined figural scoring but not necessarily by the verbal scoring because verbal elaboration is no longer used (Torrance, 1974a, 1974b). This author concurs with those who promote the necessity for inclusion of high quality of the responses in scoring for creativity (Besemer & Treffinger, 1981; Harrington et al., 1983; Hendrickson, 1985/1986). This author, then, considers the verbal TTCT divergent production scores of fluency, flexibility, and originality to be measures which reflect a more limited range of creative behaviors than the new measures of the figural TTCT streamlined procedures. The verbal divergent thinking scores, particularly verbal flexibility (see Tables 1 and 2), however, are a useful measure and indeed in this study were the most sensitive of the measures to the changes in levels of creativity associated with the OM creativity training in this study (see Tables 3, 4, 5, and Figures 3 and 4). It was surprising to note that the verbal TTCT scores, however, were lower on pretest performance for children who had prior experience in OM than for those without such OM creativity training. This finding, in conflict with the higher posttest scores for the OM-hi effort children, might be an artifact of a chance distribution of students with high pretest scores (see p. 165).

Similes

As evidenced by the intercorrelations of the data produced in this study (Tables P-1, P-12, and P-25, Appendix P), Similes shared significant correlation with all of the figural measures except for fluency, the total of creative strengths, elaboration, and resistance to premature closure and, therefore, with the figural Creativity Index, and a negative correlation with verbal originality. The regression analyses (Tables P-8, P-17, and P-20, Appendix P) confirmed the relationship of the creative strengths as a predictor variable for Similes. Similes, along with the verbal creativity measures, was responsive to the effects of differences in effort and training for creativity associated with OM (Tables 3 and 5, and Figure 4). Similes was consistently present in the cluster of creative abilities identified by the pretest measures that were associated with exposure to prior training in creativity in the enrichment program (Tables P-4 and P-5, Appendix P), in an OM team (Tables P-6 and P-7), and in the propensity to seek further creative experiences, that is, differences between children who chose to join a team and those who did not (Tables 1 and 2). As indicated in the discussion in Appendix E, the retest reliability values obtained in this study were probably an underestimate of the stability of the measure. The stability coefficient of .53 (Table F-9, Appendix F) was a relatively strong indicator of the retest reliability of Similes considering the presence of the significant intervening treatment effects. Excellent inter-rater reliability (.92-.94) was readily obtained for a random sample of 20 tests from this study (see Chapter III, Instruments). The results of this study, then, confirm previous statements of the reliability and add to needed validity information for the Similes test (Schaefer, 1971). Incorporation of Similes in the creativity battery of this study provided a valid measure for assessment of criterion validity of the new figural streamlined variables and provided the study with a measure that demonstrated excellent sensitivity to the criteria of creativity training.

Affective Instruments

How Many Ideas?

The creative self-concept scale, How Many Ideas? (Ideas) (Fishkin, 1987a, 1987b), used in this study was developed because of the need for a measure of self-concept specific to creativity that is appropriate for elementary children.

Ideas, adapted from the creative self-concept scale of Wright et al. (1975) was described in Chapter III and Appendix K.

Reliability. The reliability of Ideas was established in three ways. Internal consistency, α=.85, was demonstrated in a pilot sample of 21 graduate students. Acceptable reliability was shown in responses by children, α=.70, for 115 of the Ideas pretests completed by the children in this study. Retest reliability was assessed at r=.44 for Ideas. The method of assessing test-retest reliability for the measures in this study (see Table F-9, Appendix F) should be considered to be underestimates because of the intervening significant treatment effect during the five-month interval between tests.

Examination of the frequency distribution statistics (see Table F-10, Appendix F) revealed a remarkably normal distribution for both pre- and posttest administrations of the scale. Although four children responded with a perfect score for the pretest and two did so for the posttest, scores were distributed from 16 to 50 with only a low amount of negative skew. The internal consistency of .70 with this age group and the wide and normally shaped distribution suggest that the scale shows promising characteristics of being able to discriminate the attributes consistently in a sample of elementary-aged gifted children.

It was noted in scoring the tests that a few of the children seemed to have reversed one or another of the items on their tests in a direction that seemed highly inconsistent with most of their responses. It seems that the method of presenting reversed items may have contributed a greater source of unreliability for the children than for the adult pilot group. In her adaptation of Ideas for a study of perceptions about creativity in parents, teachers, and OM coaches, Christy (personal communication, April 16, 1988, see Appendix A) used an underline of the "seldom" and "often" anchors and adapted the instructions to assist the respondents to be more aware of the reversals. These adaptations would be likely to improve on the reliabilty of this scale.

Validity. Content validity, as discussed in Chapter III, was derived from two sources. First, the scale as developed by Fishkin (1987a, 1987b) changed the content and format of Wright's scale to only a minimal degree. Second, the content was examined by experts in creativity, in gifted education, and in reading instruction at the elementary level. These experts determined, after the adjustments to lower the reading level, that the scale was a measure assessing how children feel about themselves in relation to creativity, that is, creative self-concept, and was age-appropriate for gifted elementary children if the items were read aloud to children under third grade.

Examination of the within cell correlations of the affective variables (Table P-18, Appendix P) reveals that Ideas carried significant relationship with all variables in the affective construct. Although Ideas showed significant correlation with both measures of locus of control, $r_{I+}=.21$ and $r_{I-}=-.28$, it was more strongly correlated with the measure of general self-concept, $r_{Sears}=.64$. The prediction equation from the multiple regression analysis (Table P-23, Appendix P) confirmed the clarity of the relation of Ideas to itself and to general self-concept. The significant variables predicting posttest scores for Ideas were posttest scores for the Sears, pretest scores for Ideas, and months of experience in the enrichment program (see Table P-23, Appendix P). The clarity of the strong relationship of Ideas to general self-concept and some but remote relationship to locus of control, provides good evidence toward establishing construct validity of this new scale as a measure of specific creative self-concept.

Construct and criterion validity for the parent creative self-concept scale by Wright et al. (1975), from which Ideas was developed, were determined by comparison of performance on the creative self-concept scale, a general self-concept scale, and tests of creativity. Wright et al. found that creative self-concept carried a significant but fairly low positive correlation, $r = .24$, with general self-concept, the Tennessee Self-Concept Scale. Intercorrelations were produced among scores generated from two TTCT subtests, Unusual Uses and Incomplete Figures, and general self-concept and creative self-concept. In their sample of 80 college junior and seniors majoring in education, general self-concept did not relate significantly with any of the eight creativity variables but creative self-concept had a low correlation with verbal fluency and verbal originality, and a correlation of $r = .30$ with verbal elaboration. Multiple regression analyses confirmed the absence of a significant relationship between general self-concept and the eight creativity variables but established a relationship of creative self-concept to the creativity measures ($R = .44$, $p < .05$).

Evidence of construct validation of Ideas was provided by three processes. The first, an essential step in construct validation (Carmines & Zeller, 1979), was established by describing a theoretical network for the test. In Chapter II, the review of the literature on self-concept and creativity revealed that general self-concept scales, the Piers-Harris and the Tennessee, did not show predicted relationships to creative behavior (Fults, 1981; Jones, 1982/1983; Kolloff & Feldhusen, 1984; Wright et al., 1975). The theoretical justification for use of scales of self-concept specific to the content of the investigation was presented by examining research and theory related to self-concept specific to the content of academics (Shavelson & Bolus, 1982). The need for measuring self-concept specific to creativity was suggested by Williams (1976) and Wright et al. (1975).

The empirical relationship between Ideas as a self-concept scale was demonstrated by its clear relation to general self-concept and its minimal relation to a different area of affect, locus of control. The similarity of content and format of Ideas to the scale from which it was derived (Wright et al., 1975) et al. is applicable to Ideas. To determine the validity of this supposition, Wright's statistical procedures were replicated with the measures used in this study. Intercorrelations were generated for the 14 creative and affective variables in this study (Table P-25, Appendix P), and four multiple regression equations were computed, with predictor variables of the posttest affective means, general self-concept, creative self-concept and also the locus of control measures (Table P-26, Appendix P). The intercorrelation matrix supported the interpretation that Ideas is more closely related to a greater number and variety of the creative variables than any of the other affective variables. Ideas carried a significant positive correlation with six of the 20 possible creative comparisons, for Similes and each of the figural variables except for closure. The regression equation (Table P-26) showed Ideas to be somewhat more closely related to the creative construct than the other three affective variables. The prediction equation for Ideas was comprised of the posttest variables for general self-concept, and also for figural fluency and elaboration. The empirical findings of the intercorrelation and regression analyses indicate support for construct validity of Ideas as a measurement of self-concept and as a measurement of self-concept which relates to creative behaviors.

The intercorrelation matrix, however, also offers clear support for the theoretical argument presented in Chapter I. The creative and affective variables were shown to form two separate coherent constructs which carry correlation greater than .30 only with some of the other variables within each construct.

Evidence of criterion validity of this scale is found in its usefulness in revealing important and meaningful relationships between creativity training and affect in gifted children. Inclusion of Ideas in the affective construct was critical to revealing the important effects for obtained differences in the multivariate construct of affect by differences in time and in OM effort. Differences between pre- and posttests were found in the construct with a significant increase in creative self-concept for all children in the study (see Tables 6 and 7, and Figure 6). Although possible that these differences may merely be attributable only to maturation, it is likely that this effect reflects the potency of the creativity training present in the overall enrichment program. Moreover, Ideas was sensitive to differences between the OM groups. A significant difference was found in the multivariate construct for affect, with significant univariate differences for general self-concept and creative self-concept (see Tables 6 and 8, and Figure 7). The differences were found between OM effort groups in posttest creative self-concept scores adjusted by MANCOVA procedures for differences in pretest scores, sex, and amount of prior experience in the enrichment program or on an OM team. This analysis of adjusted posttest scores indicated that students who exerted a high amount of effort in their OM team CPS work were higher in creative self-concept than children who were low in effort on an OM team or who were not in OM. It should be noted, however, that the results of the MANCOVA analysis, although significant, bore a very low strength of association. That is, estimates of the strength of association indicated that the effect of OM on creative self-concept was able to account for .02 of the variance in Ideas. As stated in the beginning of this thesis, there is a lack of significant findings of relationship of CPS or even of creativity training to self-concept in the research reviewed by this author. Studies referred to by Feldhusen and Clinkenbeard (1986) and by Parnes (1972) suggest that some changes in aspects of personality have been found to be associated with training in creativity and/or CPS but these findings are less consistent than results associated with creative measures. Given the paucity of research evidence and the low level of association in this significant finding, it may be that changes in self-concept, when they occur in association with creativity training, may be effects of a low order of power. Many significant findings of research in the behavioral sciences are, however, of a low order of power, in fact, the vast majority of studies "do not account for more than 10% of the variance" (Linton & Gallo, 1975, p. 331).

Ideas appears, then, to be a measure that can be fruitfully used in an area of research where appropriate measures had been lacking. It has acceptable reliability for an instrument whose primary use is for research purposes. Validity evidence for Ideas as a scale appropriate for assessing changes in creative self-concept in children is promising. Moreover, the scale is administered quickly and was sensitive to appropriate and important effects in this study.

Intellectual Achievement Responsibility Scale

The IAR was selected and used for this study because of relevant research cited by Crandall (1978) which indicated its validity with task persistence, creativity, and academic grades and also by McKenna (1981) for effort and achievement test scores. As described in Chapter III, previous studies of reliability obtained statistics of consistency that were less than satisfactory. Reliablity coefficients obtained on the pretest data for 102 of the children in this study were at I+=.48 and I-=.50 for Spearman-Brown procedures, corrected for attenuation, and I+=-.29 and I-=-.13 for coefficient alpha. Retest reliability in this study was .52 for both scales (see Table F-9, Appendix F) but as previously stated, the retest reliability coefficients obtained for measures in this study are very likely to be underestimated. In addition, the retest reliability computation for the IAR included the scores of eight children whose pre- or posttest protocols were only 85 to 91% complete (see procedure for missing data). In at least two of these cases, it was clear that it was the child's intention to respond by leaving the items blank. Nevertheless, the reliability coefficients for the children in this sample indicate that it is possible that as much as 72% of the variance in I+ or I-could be error variance and, therefore, the IAR would be unlikely to yield meaningful results in this sample.

Examination of the frequency distribution data (Table F-10, Appendix F) revealed that the distribution for I+ bore a marked negative skewness, particularly for posttest I+ (-1.09). These results indicated that the test did not discriminate for a full 33% of the sample who missed no item or only one item of the 17 for a perfect score. The scores were lower for I- with a near-normal distribution for pretest and only a moderate negative skew for posttest I-. Inspection of the mean values for the items located five items for the I+ scale which resulted in a mean score of less than .20 positive responses to I+ (Items 2, 5, 6, 17, and 24) and three items in which there was clearly a ceiling, a mean of greater than or equal to .90 for items 1, 16, and 31 (see Appendix L). The scale was more coherent for I-, with the range of item means distributed somewhat more closely, from .15 to .87. The four items which yielded a mean value of less than .25 were items 7, 10, 18, and 26. It is possible that some of these items will not discriminate locus of control accurately in a gifted population. For example, the item pair of I- #10 and I+ #24 reads for #10, "If a boy or girl tells you that you are dumb, is it more likely that they say that A. because they are mad at you, or B. because what you did really wasn't very bright?" A gifted child may perhaps be less likely to be called "dumb" than the average child, but when such name calling occurs, it indeed may more often be irrelevant to that child's actions and more likely based on jealously.

No significant effects were noted for I+ in this study other than a positive relationship with two of the three other affective variables. A significant positive correlation for posttest I+ was shown with pretest and I+, with the Sears general self-concept, and with a weak but significant relationship with creative self-concept (Ideas) (see Tables P-18 and P-25, Appendix P). Locus of control for negative events, I-, showed relation of the posttest scores with pretest scores and a weak relationship with Ideas (see Table P-25). There were other significant findings, however, for I-: an overall increase in posttest scores across all OM groups thus showing a possible effect of increase in locus of control associated with the enrichment program itself. There was a significant correlation of I- with figural elaboration and with the total of creative strengths (see Table P-25), two of the figural variables which could be interpreted as having components of motivation or strength of creative energy. "The elaboration score has consistently correlated higher with measures of school achievement than any of the other measures of figural creativity" (Torrance, 1979, p. 65).

McKenna (1981) conducted a study examining the relationship of the IAR to intelligence, achievement, and to effort in academic work of gifted children. She found both the I+ and I- scores to bear a significant correlation with grades of effort in reading but not with grades of achievement in reading. The relationship of IAR scores to effort was found for fifth grade but not for the third grade children in her sample.

The present data, showing lower scores for I- than for I+, are consistent with the literature on the IAR and on attribution of causes indicating that responsibility is less readily assumed for failures than successes (Bennett, 1984; Crandall, 1978). A comparison of the grade level means with the norms in Crandall, Katkovsky and Crandall (1965) indicates that, in the present sample of gifted children, there is a higher acceptance of responsibility for successes but a lower acceptance for failures than in the norm group. Internality for "failure" experiences was increased (see Tables 6 and 7, and Figure 5), but posttest scores for I- were below those for the respective grade level norm groups for all OM conditions.

Sears Inventory

The results of this study showed that the Sears was remarkably consistent in measurement from pre- to posttest , r=.69 (Table F-9, Appendix F), in comparison to the other measures of this study. As discussed in Appendix E and in this chapter under reliability of the TTCT, the retest reliability values obtained in this study should be considered as underestimates of reliability because of the significant intervening treatment conditions present during the time interval between pre- and posttseting. The retest reliability in this study actually resulted in a higher retest reliability than the r=.50 obtained by Sears (1972) where there were no known treatment conditions present. The Sears is least like the other measures in the study which have been known to bear a significant relationship with intelligence and/or achievement (TTCT, Torrance, 1974b; IAR, McKenna, 1981). Because of the relation of the creativity measures and the IAR, and perhaps also Ideas, to intelligence and/or achievement, those measures might be expected to be more limited in their range of distribution, or, at least, less normally distributed than the Sears. Inspection of the frequency distribution statistics of the measures (Table F-10, Appendix F) indicates that there was a large spread of scores and negative skewness, pretest=-.63 and posttest=-.85. Although skewed in distribution, the histograms revealed a more normal shape in distribution of the Sears than that of the TTCT verbal distributions.

The Sears shared relationship with positive attribution of responsibility (I+) and with creative self-concept (Tables P-18 and P-23, Appendix P). As predicted, the Sears was the most stable of the affective measures when examining the measures for change over time (Table 7). However, analysis of the posttest scores adjusted for pretests, for experience in enrichment and on an OM team, and for sex revealed a significant effect on the Sears Inventory by differences in effort of the OM groups (Tables 6 and 8). The effect indicated that there was a decrease in level of general self-concept of children in the low effort group of OM. It must be noted, however, that the strength of association of this effect is at a fairly low level, N^2=.05, thus indicating that the effect of OM explains only 5% of the total variance in the Sears dependent variable. Although at first glance surprising, lowered self-concept for children who were perceived by themselves and their teammates to have exerted little effort on the team is a likely, but unfortunate, concomitant effect, when people do not do an adequate job of engaging themselves in work to be done. The effect of competition, even when participating in the competition with teammates, and/or perception of competition were not assessed in this study as distinguishable from amount of effort. Some authors have found that many children report that competition has a meaning in their history of experiences that is debilitative (Roweton, 1982).

It has been observed that gifted children are often in need of learning skills in interdependence; for example, in how to work cooperatively in a group, in expressing their needs, and in facilitating the functioning of others within the group (Torrance, 1985; Webb et al., 1982) (see p. 3). Therefore, it is not surprising that this study found working on an OM CPS group to be associated with lowered self-concept by some of the gifted children. It has been found that achievers in an academic setting have higher self-concepts on intellectual and school status subscales and on total self-concept, and higher internal locus of control for negative events (higher IAR I- scores) (Kanoy et al., 1980). Likewise, students who did not exert effort in an OM team were not achievers in the OM setting and thus might feel less positive toward themselves with respect to those aspects of self-concept. This interpretation is commensurate with an increase in creative self-concept in children who were not in OM and in OM-hi, and no increase in children who were OM-lo (see Tables 6 and 7, and Figure 8).

Unlike prior research on self-concept associated with training in creativity utilizing the Piers-Harris Self-Concept Scale (Fults, 1981; Kolloff & Feldhusen, 1984; Piers, 1977), this study found differences in general self-concept associated with creativity training in the form of the amount of effort exerted on work on an OM team. However, because the effect accounted for only a relatively small amount of the variance in self-concept, and this study was a quasi-experimental study, more definitive answers are still needed. Moreover, it remains to be determined whether the results are accounted for primarily in the questions on the Sears which address the content of creative self-concept or in the questions which tap the more general areas of self-concept.

The results of this study, then, confirm earlier results that the Sears Inventory is a reliable instrument. The findings also seem to confirm Whitmore's statement that the Sears is a sensitive instrument for assessment of self-concept (personal communication, August 15, 1985, see Appendix A). It was an appropriate measure for use with this sample of gifted children.

Relation of this Study to Some General Issues in Creativity Research

Effort and Creativity

Use of effort in OM CPS provided the key to obtaining interpretable results in this study. Examination of those significant findings which differentiated the OM groups clearly shows that the main differences were between children who were in OM high effort versus low effort rather than between children who were in OM and were not in OM. As may be seen by examining Figures 3, 4, and 6, with the exception of Similes, there was a greater difference between the OM-hi and OM-lo groups than between OM-hi and non-OM for all of the major variables which obtained a significant effect for OM: mean verbal TTCT, Ideas, and Sears.

On Similes, OM-hi was different from both OM-lo and non-OM. If the design of this study had simply compared students who were in OM with those who were not in OM, it is most likely that nonsignificant findings would have been obtained as in the case of other studies of affect and creativity training (Fults, 1981; Jones, 1982/1983; Kolloff & Feldhusen, 1984) and some of the studies which did not show changes in creative thinking (Tadlock, 1981).

It must be noted that the non-OM group served as a control group in this study but was not a true non-experimental control group. That is all students in the study including those who were in the non-OM group were engaged in creative and affective training activities in their enrichment classes (see Chapter III and Appendix D and N). Considerable exposure to CPS processes of brainstorming and interactive team process was included in the students' class time work on problems of the Future Problem Solving program (Crabbe, 1985) and other in-class activities.

The results of this study, then, reveal that effort, which was previously unexplored in creativity research as a method for differentiating groups of subjects was a valid variable to divide the groups in this study. The use of effort created an important blocking variable which accounted for effects related to creativity training in a meaningful way. It is likely that student effort on a team may be related to intrinsic motivation. The variable identified as effort resulted from definition of an observable, measurable, and reliable operational definition of defining the effort produced by an individual. This definition of effort was shown in Chapter II to have theoretical foundation (Franks & Dolan, 1982; Osborn, 1957; Renzulli, 1978; Torrance, 1979) and have meaningful relation to the sparse relevant research (Bennett, 1984; Blumenfield et al., 1986; Friedman et al., 1984; McKenna, 1981).

Effects of Social Influences and Creativity

The effects of social influences on creativity as discussed in Chapter II are likely to be very relevant to the findings of this study. Research on the effects of creativity training in social groups have yielded inconclusive results. Some authors concluded that work in groups on creative endeavors is harmful to creativity (Abramson, 1976; Amabile, 1983). Others assert that creative abilities are enhanced when participants can work together in a group (Osborn, 1957; Parnes, 1972, 1975; Roweton, 1982; Torrance, 1972a). A closer examination of some of the studies reveals that the comparisons of groups versus individuals were frequently contaminated by factors such as "real" groups versus non-participating groups (Abramson, 1976; Taylor, cited by Stein, 1974-1975) or audience effects (Cottrell, 1968). Two findings seem very clear; creative behavior is a very complex behavior and creative performance is heavily influenced by motivational effects (Torrance, 1979; 1987) (see discussion below under creativity and motivation). Perception of judgment, or even likelihood of judgment of one's creative efforts, also appears to lead to less creative behavior (Amabile, 1983; Parnes, 1972).

If, as Amabile (1983) stated, the presence of others leads to an increase in arousal and perhaps to expectation of evaluation, then a decrease in creative performance would be expected. Others have expressed similar concerns that it is important in group CPS to establish a non-evaluative climate, that is, to exercise the principle of deferring judgment during brainstorming, in order to actualize creative potential (Bull & Fishkin, 1987; Torrance, 1979) for when people are made anxious or fearful of expressing their ideas or feel their ideas are restricted, then less creative behavior will result (Parnes, 1972).

The impact of this study bears on the critical "if . . . perhaps . . . then" quality of Amabile's position (1983, p. 142). In this study, often within the same team, some of the children showed much greater creativity as measured by Similes, no decrease on verbal TTCT, and an increase in creative self-concept, while others suffered significant declines in verbal creative performance, as measured by the TTCT, and in general self-concept, and showed almost no increase in Similes and in creative self-concept (see Tables 3, 4, 5, 6, 7, and 8; Figures 3, 4, 6). These results clearly imply that individual differences in effort invested in the creative enterprise have significant bearing on the "if . . . perhaps . . . then" effects of evaluation expections and impact on creativity in a given group setting which may be very beneficial to the development of creativity and creative self-concept of some of its members.

This study can offer no definitive answers about the relationships of social influences on creative behavior. It is possible that there may be a confounding effect of results of competition, of perceptions of judgment by other members of their teams, and of individual effort. Some additional analyses of the data may shed further light on possibly untangling these relationships. The use of effort as an independent variable in this study has helped to shed some additional light on some of the conflicting findings of effects of social groups and creativity. The children who were high in effort in their OM teamwork showed increases in creativity or at least were able to withstand the very potent deleterious effects on motivation to be creative producers on the verbal TTCT. The high effort children and those not in OM showed an increase in creative self-concept. The increase in creative self-concept for those who were not in OM can probably be interpreted as reflecting the training effects of the overall enrichment program in which all groups in the study participated. Therefore, these findings are supportive of the statements in Chapter II, that teamwork which is experienced as collaborative provides a situation which facilitates growth in creative abilities and may contribute to creative self-concept.

Social Influences and Affect

The preceding section describes some of the mechanisms which may have contributed to the effects noted for the affective construct. It is possible that there may be a confounding and possibly interactive effect of winning or losing in competition, of perception of evaluation by others, and of amount of effort exerted in the creative enterprise of the OM team's work.

The rather disturbing finding of lowered general self-concept in children who were low in effort must be addressed. Two aspects must be considered. It is possible that it was a chance effect, because the strength of association revealed that although a significant finding, the decline for OM-lo students only accounted for .01 of the total variance of general self-concept in the children of this study. If it is a "real" finding and replicable, then some recommendations about nurturing creativity and nurturing the individuals within the team who may be reticent team members would be in order.

There is some conflict in the literature on gifted children and effect of various programs on the self-concept of gifted children. Some authors have found that differences in type of programming can account for changes in general self-concept (Coleman & Fults, 1982; Janos, et al., 1985; Maddox, Scheiber & Bass, 1982; McQuilkin, 1980/1981). Fults (1981) found lower self-concept in gifted children who attended pull-out gifted programs than in high-achieving children who were in mainstream classes. Because their comparison group is limited to the very bright, the gifted may feel less adequate about themselves (Coleman & Fults, 1982). In a follow-up study, Coleman and Fults (1982) found that although there was a lower self-concept on the Piers-Harris for the children who were in the homogeneous pull-out gifted class than for the children who were mainstreamed, both groups had a higher mean self-concept than children in Piers's age comparison group. Moreover, the effect of lowered self-concept was temporary because the effect disappeared for those children who had graduated from the program. Coleman and Fults considered the negative effects on self-concept to be of a transitive nature, merely reflecting the effects of use of the "gifted comparison group for evaluative purposes" (p. 119). Others found that the type of grouping was not really relevant; whether the children felt encouraged and nurtured in their class environment was more critical than the type of class grouping. A well-designed study by Kolloff and Feldhusen (1984) confirmed that there were no differences in self-concept between gifted children in pull-out programs and those who were not on either the Piers-Harris scale or the ME Scale by Feldhusen and Kolloff, a self-concept scale designed for gifted students. Their results (of no differences between the groups in self-concept) were interpreted to indicate that "enrichment programs for the gifted do not change self-concept either positively or negatively" (p. 56).

The lowered general self-concept in OM low effort children, then, could be of a transitive nature reflecting work within a comparison group where perhaps the others who had exerted more effort were judged to be more competent. The results could also reflect the sensitivity of the Sears instrument to a comparison group in which creative abilities constitute part of one's general self-concept (see Appendix M).

Motivation and Creativity

Tests of creative thinking abilities, including the TTCT, have been shown to be very sensitive to variations in a range of factors which can affect motivation to produce creative energy such as testing environment, administration of the battery, presence and kind of warm-up prior to the test (Torrance, 1972a, 1979, 1987). Because of the sensitivity of these tests to motivational influences, this author agrees with Khatena (1982) that it is inappropriate to think of scores on a test of creative thinking as some absolute measure of that person's native creativity. Without motivation to behave creatively, even though the individual may possess the ability and the skill to behave creatively, it is unlikely that creative behavior will emerge (Torrance, 1979). This author considers the sensitivity of creativity tests to motivational and environmental variations of test administration, although a source of potential unreliability of test scores, to contribute to the validity of the tests.

Two earlier studies (Mohan, 1970, cited by Torrance, 1972a; Elkind, et al., 1970) are discussed in Appendix E in relation to the likely impact of relevant motivational factors on TTCT scores for the children in this sample. Mohan found that highly creative children show greater responsiveness than children who are low in creativity to variations in richness of cues in the testing environment. A reasonable implication from the Mohan study is that highly creative children would be likely to be highly responsive to motivational influences affecting the testing session. As discussed in Chapter IV, pp. 141-145, the children in this study were found to be remarkably creative (see Tables 1 and P-2, Appendix P).

The context effect, the effect of type of ongoing activity which is interrupted by a testing battery, was shown to have a potent impact on scores on creativity tests (Elkind et al., 1970). Their study is summarized in Appendix E and briefly discussed above under effects of time on the creative construct. This researcher concludes that the decline in verbal TTCT and figural elaboration scores is primarily attributable to lessened motivation of highly creative children to expend their full potential of creative energy because of a potent context effect. It is also likely that motivation for the verbal TTCT may have been adversely affected for the third grade children because of their recently concluded extensive achievement testing (see p. 125). It is also attributable to an effect of regression of exceptionally high pretest scores to decrease in value toward the mean.

Intelligence and Creativity

A significant, but low, correlation is observed between tests of intelligence and creativity (Getzels & Jackson, 1962). Others report little or no relationship between intelligence scores and creativity measures (Wallach & Wing, 1969). The relationship for the TTCT battery lies primarily in the verbal rather than the figural tests (Shaffer-Zamari, 1983; Torrance, 1974b). Davis and Rimm (1985) considered MacKinnon's threshold view of the relationship between creativity and intelligence to offer a "good resolution to the creativity-intelligence issue" (p. 211). That is, there is a moderately good relationship between creativity and intelligence when the full range of intelligence is included. However, when examining only the higher range of "intelligence--namely, above a threshold IQ score of about 120--there is no relationship at all" (p. 212).

However, the remarkably high pretest scores on the verbal TTCT and scores on Similes show that the children in this sample should be characterized as creative as well as gifted. The mean verbal TTCT score, 132.14, which was obtained on only six of the seven activities, reflected mean performance on the creativity tests that were one and one-half standard deviations above the mean of the norm groups for the TTCT. In this sample of gifted children, the IQ ranged from 130, almost two standard deviations above the mean, to the ceilings on the tests of individual intelligence. As a test of general intelligence, the Stanford-Binet Intelligence Scale provides a strong sampling of tasks of verbal intelligence for those children who are tested primarily at the ranges of middle childhood and adolescent age tasks (Clark, 1983). The children in this sample, then, were mentally gifted, and many showed particular strength areas of verbal intelligence such as abstract and verbal reasoning, memory for verbal material, and vocabulary. Because training for creativity can increase the person's abilities in creative thinking (Davis & Rimm, 1985; Feldhusen & Clinkenbeard, 1986;

Rose & Lin, 1984; Torrance & Safter, 1986), because scores on tests of creativity from children who are highly creative are responsive to environmental conditions (Mohan, 1970, cited in Torrance, 1972a), and because training in creativity is recommended as an important part of gifted education (Clark, 1983; Davis & Rimm, 1985; Maker, 1982), then it would seem likely that there might be a correlation between intelligence and creativity once the children have received service in a gifted program which provides training in creativity. The obtained high mean scores for the verbal TTCT, figural elaboration, and Similes (see Tables 1, P-2 and discussion in Chapter IV under adjustment of ceiling for TTCT scores), then, suggest that a moderate relationship between intelligence and verbal creativity could exist in this sample. These data suggest that perhaps the threshold concept of the creativity-intelligence issue (Davis & Rimm, 1985) ought to be modified so that the threshold concept may only apply for those who have not been exposed to training in creativity. Perhaps the correlation between IQ and creativity may occur over the threshold of IQ of 120 as well as throughout the range of IQ in those who have been exposed to creativity training. Further research is needed to address these issues satisfactorily.

Grade Level and Creativity

A significant amount of the variance in creativity in this study is attributable to differences in grade level and to the interaction of grade and changes over time. Estimates of the strength of association of the amount of variance in the effect of differences in grade on the multivariate construct of creativity showed that the differences in grade were accounted for most strongly by the measures of verbal flexibility, verbal originality, and figural closure. The greatest differences were found between Grade 3 and Grade 4. These results are consistent with earlier researchers (Torrance, 1974b; Torrance & Safter, 1986; Williams, 1976). The fourth grade "slump" in creativity has been observed to occur in cross-cultural studies (Torrance, 1974b), as well as in numerous studies of creative behavior in American children. It must noted that it is possible that some indeterminate portion of the lowered scores for third graders is due to an artifact of this study. That is, the extensive spring achievement testing for third grade students may have contributed to their lowered motivation to expend creative energy on the verbal TTCT (see p. 125).

The literature surveyed by the author does not address the possibility that some of these differences in performance between third and fourth grade could, in part, be an artifact of differences between individual and group administration of the TTCT. The gifted chidren in this study were also a highly creative sample and, as highly creative children, might be likely to have greater potentially different levels of response on a creativity test because their responsiveness to variations in motivational conditions. The presence of an individual adult examiner who is writing down all of a child's responses usually is likely to be viewed as a motivating condition, and perhaps, particularly so by gifted children. This study, however, was not designed to control or assess the effects of differences in grade level on creativity. Therefore, resolution of possible issues of this fourth grade "slump" must remain with other investigators.

Selection of Appropriate Instruments

Creative Measures

As discussed in the creative measures above, use of the composite scores served to clarify the results. However, when the standard scores of the figural TTCT were used to compare the subscale measures, these scores were less reliable than the raw score measures for closure, titles, and possibly elaboration and total of creative strengths. The raw scores were more reliable for these subscale measures because of the distortion effects produced by the conversion to standard scores upon the limited spread of these scales (see Tables F-6 and F-7,

Appendix F). As discussed in Appendix E and in this chapter, the composite scores, then, may be more reliable and useful for purposes of identification than the scores of the individual scales.

The results of this study lend support to the common practice of selecting only some of the activities of the verbal TTCT. In this study, where time was taken from classes to administer the affective measures as well as the creative measures, use of the full battery of the six activities contributed to a motivational response set that was counterproductive to expenditure of creative energy (see Footnote 2). In retrospect, and in considering Torrance's recommendation for priority of selecting test activities, in this study, use of only two activities, Product Improvement and Just Suppose, alone might have produced a more valid measure of verbal creativity. The full figural battery appeared to be sufficiently motivating for all three activities to be used. It is possible, considering the limited range of scores of abstractness of titles, that if Activity 1, Picture Construction, were to be deleted from the battery that the titles measure might lose too much of its reliability to be a useful predictor of the criterion of interest. Similes showed significant and meaningful differences in relation to several of the effects of interest in this study. Based on the findings in this study, this 15-minute test continues to be a reliable, valid, and useful test of creative behaviors.

Affective Measures

The creative self-concept scale (Ideas) (Fishkin, 1987a, 1987b) showed adequate reliability and promising validity in its sensitivity to the effects in this study as shown by the support for hypotheses 11 and 12. The measure is brief, is easily administered, and seemed to hold good face validity for the children.

Footnote 2: Activity 6, "Unusual Questions" will be deleted from the forthcoming verbal streamlined procedures (Torrance & Ball, in press).

In retrospect, the IAR may not have been the best measure of locus of control for this study. The scale is designed to be specific to tapping school-related achievement expectancies, and the research shows it often reflects effects related to achievement in school (Crandall, 1978; McKenna, 1981). It is possible that the effort involved in OM related tasks is sufficiently distinct in character from effort in school-related tasks that the IAR was not able to distinguish meaningfully between the OM groups. This explanation makes some sense in light of the effect produced on I- for time, showing an increase in posttest as compared to pretest scores. Therefore, the increase in posttest scores may reflect an increase in locus of control associated with the training to increase self-directed behaviors in all of the children in the program.

Two other qualities of the IAR, however, may contribute to the lack of significant findings on the IAR in relation to OM and even to experience in the program. Scores on I+ were quite high; the distribution was negatively skewed with several children obtaining a perfect or near perfect score. The items may not offer sufficient spread to discriminate I+ in a sample of gifted elementary children. From comments from some of the children, the use of an audio-cassette tape method for administering the instructions and reading the items appeared to reduce the face validity of this measure. However, since an overall increase in I- occurred for this study, the face validity is assumed to have been adequate.

The most important problem with the IAR was the reliability of the measure. The split-half method test of reliability of the IAR, as cited in Chapter III, is consistently lower than .60. In this study, the split-half reliability of I- was only .50, and coefficient alpha reliability was even lower at -.13. Coefficient alpha reliability is the most generally useful method of determining reliability and is superior to methods based on subdivision of the test which often results in an overestimation of the reliability of the measure (Carmines & Zeller, 1979; Nunnally, 1970).

However, the problems with the IAR might, in part, lie in that measures of attribution for success or failure experiences which are common for many children do not relate appropriately to the experiences of gifted children. In her discussion of the failure of measures of causal attribution to relate to observed high or low classroom achievement in mentally gifted children, Bennett (1984) noted that some of the questions of the Sydney Attribution Scale might not have been meaningful to the highly gifted children in her sample. Likewise, in the present study, it was noted that at least three questions could perhaps have been more meaningful to the gifted children in the study if the questions had been recoded.

> Perhaps, total failure is an unfamiliar situation for {gifted} students who may have only experienced . . . 'less success' than expected. Thus, to more clearly understand gifted students' perceptions of failure it may be necessary to use other kinds of measures or to describe failure in terms which are more relevant to the experiences of gifted students (Bennett, p. 75).

The Sears Inventory, as discussed above, showed excellent reliability in this study. Further study of its subscales for self-concept toward divergent thinking is planned. The measure was, as predicted, relatively stable over time. Unlike studies using the Piers-Harris, the measure was sensitive to differences between the creativity training groups. The Sears revealed a lowered self-concept in those children who exerted low effort in OM. The measure appears to have been an appropriate choice for studying self-concept in children where the criterion treatment effect relates to creativity training.

Conclusions

Because the study is a quasi-experimental study, statements about causality of treatment effects are not appropriate. Because this study did not use a randomly selected sample for the study, cause-effect statements cannot be attributed to the OM treatment effect. However, the presence of two levels of comparison groups, non-OM and OM-lo, and the control of pretest scores for comparison purposes provide this study with some of the important rigorous qualities normally attributed to true experimental research. Therefore, any obtained treatment effects in this discussion will be considered as suggestive of possible causality which would be fruitful areas for possible confirmation by further research (DeLisle, 1985; Isaac & Michael, 1981).

Generalizability

These findings should be generalizable to elementary students on OM CPS teams and possibly to other gifted populations in schools which use pull-out programs and/or other programs which emphasize CPS. The generalizability of the findings is probably not appropriate for programs for gifted children where the goals of the curriculum are centered primarily on academics and where little team CPS is offered.

The internal validity of this study appears to be threatened by the effect of prior testing on the scores of the subsequent testing (Isaac & Michael, 1981). In this study, most students were tested during their one-half day a week enrichment program. Total time of testing over the year accounted for eight hours, over 10% of their total time in a program in which there is normally no testing. During posttesting, several students complained about losing time from preferred learning activities in order to take tests. It must also be noted that this study was conducted with a highly creative population; overall mean scores in verbal originality were 126.93, and overall mean scores in figural originality were 120.14.

With pretest mean scores greater than one standard deviation above that of the normal population, some posttest regression of the scores toward the mean would be expected. There was a significant decrease, then, in verbal TTCT performance which appears to be in part attributable to the phenomena of extremely high scores regressing toward the mean. However, this author considers the low verbal posttest scores to be best explained by the presence of a context effect, the detrimental impact on motivation for testing because the activities that were being interrupted were of high intrinsic interest to the children (Elkind et al., 1970; see discussion above and in Appendix E).

Several additional uncontrolled variables impacted this study to contribute to its not being a clean study. The children in the study were unequally distributed among the three OM groups with respect to grade, sex, amount of previous experience in the enrichment program, amount of prior experience on an OM team and/or schools attended. When the school variable was simplified to three levels to identify the three different teachers, multiple regression analyses revealed that each one of these uncontrolled variables had a significant impact on at least one of the dependent variables.

The unequal distribution of variables caused there to be too few subjects in the cells of the 3x2x2 MANOVA analyses to provide an adequate test of the hypotheses for interaction of the experience and OM variables. Results of MANOVA and of multiple regression analyses may be suspect where there are too few cases for the number of dependent variables (Tabachnick & Fidell, 1983). Two transformations were, therefore, imposed on the MANOVA procedures to examine the important treatment effects with greater confidence.

Scores of the 47 children who earned a standard score on one or more of the TTCT scales at more than three standard deviations above the mean were set at less than or equal to 160 in order to reduce the effects of these extreme scores.

The change in standard deviations of these cells reflected a more normal distribution with scores set at less than or equal to 160. Reduction of the number of creative dependent variables from 10 to 3 was accomplished by using the figural Creativity Index and the mean verbal TTCT as meaningful composite scores for the TTCT, along with the existing raw score for Similes.

Summary of Major Findings

The Major Hypothesis: Was Work on an Odyssey of the Mind Team an Effective Training Experience?

The primary hypothesis of the study, hypotheses #1 for creativity and #2 for affect, proved to show significance. That is, when the data were "cleaned up", when some of the excess variability contributing to the error variance in this study was controlled, the pattern of the effect emerged. There was a significant interaction between the OM effort groups and time on the creative construct (see Tables 3 and 4, and Figure 3). These differences showed a decrease over time on the verbal TTCT scores for the non-OM and low effort groups but no decrease for the high effort group. The statistical use of MANCOVA procedures to further control the important error variables of grade and sex clarified the relationships between the OM groups. As is seen in Tables 3 and 5 and in Figure 4, when the posttest scores were adjusted for differences in pretest scores, grade level, and sex, the OM high effort group was higher than either of the other two groups on both verbal creativity measures, on Similes and on the averaged standard scores for all three verbal TTCT scales of fluency, flexibility, and originality. These results suggest that children who were average or high in effort in OM may have benefited from working on the OM team in their verbal creativity. Children who exerted less effort on their teams than the others or who were not on an OM team did not gain as much in verbal creativity as measured by Similes and showed a significant decrease in the verbal TTCT scores as explained above under the context effect and regression toward the mean.

A tentative conclusion from the results analysis of the experience variables is that the creativity training experienced by the children who were on OM teams and/or were in the enrichment program for 13 or more months is associated with an increase in figural creativity and in Similes. On the basis of these findings, one would expect a concomitant increase in at least the creative strengths and closure to be associated with training in OM, that is, to show an effect on the OM x Time interaction. This was not the case, however; neither the mixed model MANOVAs nor a MANCOVA (see Appendix P, Table P-24) yielded significant findings for any of the figural TTCT variables for the OM x Time interaction. Moreover, the regression analyses (Table P-8, Appendix P) showed that neither of the experience variables was an important predictor in the equation for posttest performance for any of the ten creative dependent variables; therefore, these experience variables were not included as covariates for the MANCOVA. These findings appear to offer some conflicting evidence, then, on the effects of OM CPS training on the figural creative variables.

Of the creative dependent variables which were associated with the effects of previous experience, only Similes and verbal flexibility were shown to relate to the treatment effects of OM in this study. Similes showed the greater increases (see Tables 3 and 5, and Figure 4) in the non-OM and OM-hi groups. The OM-lo group, although receiving the experience of creativity training in the enrichment program, did not benefit to the extent of the other two groups. Verbal flexibility was surprisingly associated with lower pretest scores for children with experience in creativity training than for those who were inexperienced. However, of the verbal TTCT measures, verbal flexibility was the one measure which most strongly showed the effects of OM training on the posttest scores as a positive increase for those students who were high in effort on an OM team, as may be seen in Tables 1 and 2. The significant univariate effect for OM x Time was shown for verbal flexibility in the absence of an associated global effect for the construct. However, when the analysis was examined more clearly without the confusing effects contributed by measures which were too strongly correlated (see Table P-1, Appendix P) for effective multivariate analyses (Tabachnick & Fidell, 1983), the significant effect of OM and time on the mean verbal scores (Table 4) supported the findings of Table 2.

For the affective construct, the mixed model MANOVA did not support the primary hypothesis, but when the effects were analyzed by MANCOVA procedures, hypothesis #2 was supported. As is shown in Tables 6 and 8 and Figure 7, the differences are associated with a higher creative self-concept for the OM-hi effort group than either of the other two groups and a lower general self-concept for the OM-lo effort group than the non-OM group. The differences between the OM-hi effort group and non-OM group in general self-concept were not significant. These results supported hypotheses 11 and 12.

As noted in Chapter IV in reporting the results of the affective analyses, although reaching levels of significance at $p < .05$, none of the univariate results showed a strong association with those effects. Of the univariate affective findings, the strongest of these weak effects was found in the MANCOVA test for differences between OM groups on Ideas. The N^2 strength of association test revealed that .07 of the variance in creative self-concept was accounted for by differences between the OM groups. These results indicate a significant, but weak effect on the construct of affect which was accounted for by differences in effort on an OM team. These differences suggest that children who were high in effort in OM, when compared with children who exerted less effort on their team, felt themselves to be more creative persons. These results made sense with the review of literature presented in Chapter I and II. It appears that the amount of effort children invest in creative enterprise, at least in work on team OM CPS, can contribute to their concept of self in relation to their creativity. The results also seem to suggest that children who exert little effort on a team feel less confident in themselves in general than children who were in the enrichment program but who were not on a team.

What can we glean from these results, then, to shed some further light on the relationship of affect to training in team OM CPS? Two conclusions can be presented with confidence. First, changes in affect can indeed be identified which are associated with training in creativity when the measures are appropriate. In this study, two measures of self-concept--one of general self-concept (Sears) and one of specific self-concept--were chosen or developed to insure that the items reflect the content of creativity. However, these differences appear to be of a weak effect size. Second, changes in affect can be identified when the groups are examined in relation to how much effort the individuals invested in their work.

It would, therefore, be more appropriate to rephrase the problem: What is the relationship of affect to effort invested in work on an OM CPS team? When the problem is clarified to account for individual differences in effort in OM team CPS, then we can state that the OM experience appears to be beneficial in regard to creative self-concept, and, as seen in the discussion above, in regard to verbal creativity as measured by Similes and the verbal tasks of the TTCT for those children who were high in their effort on their OM teams. However, the results also indicate a finding of concern: children who were low in their effort on their team showed no difference in their measured creativity and essentially no difference in creative self-concept from children who did not work on a team, but did show a significantly lower general self-concept when compared to the control group of children who were also in the enrichment program but not on the team. Effort in OM, then, as an independent variable in this study successfuly accounts for some of the conflicting findings identified in the review of the literature (Amabile, 1983; Bolen & Torrance, 1978; Osborn, 1957; Roweton, 1982).

If the results of this study are replicable, what conclusions then can we say about the effectiveness of OM? This study clearly shows that we cannot account for differences in children if we look at OM participation versus non-OM participation. Individual differences in the amount of effort, or perhaps even amount of creative energy (Torrance, 1979) that the participants invest in their creative enterprise do account for a significant amount of obtained differences in creativity, creative self-concept, and in general self-concept. We can say that when individuals work hard and they put effort into their OM work, they benefit from the experience in creativity and in creative self-concept. The findings also suggest, however, that children who put little effort into their OM effort do not experience the beneficial effects in creativity and in creative self-concept and show a decrease in general self-concept.

How likely are the affective results to be valid and generalizable? In light of the context effect which probably depressed creative TTCT scores, the differences obtained in the affective measures can be presented with greater confidence than warranted by the level of the strength of association tests for the affective findings. The sequence of presenting the tests in classes was Similes and the figural TTCT on the first day of testing, verbal TTCT on another day of class, and affective measures one to two weeks later. In light of the depressed verbal TTCT scores, then, the affective findings are more impressive: significantly higher posttest than pretest scores in creative self-concept (Ideas) and in locus of control for negative events (I-) (see Tables 6 and 7 and Figure 6).

It is possible that there may be some limitations to generalizing these results to OM teams elsewhere because the OM coaches in this district, unlike those in many other districts, were not teachers. The volunteer parent coaches attended a two-hour coaches training workshop conducted by two of the three enrichment teachers who served as tournament directors and contact persons to coordinate the local OM program. Perhaps coaches who have more extensive training in OM and in CPS techniques and/or who are teacher coaches rather than parent volunteer coaches may bring to their coaching more skill which could better facilitate the effort and associated general self-concept in children who in this study were identified as low effort. Although the training emphasized focusing on the process of CPS and benefits of team participation for all of the children on the teams, rather than focusing on winning the competition, we are dealing with competitive children and competitive parents. The gifted children in this study were from a moderately high socioeconomic community where achievements are prized. On the day of OM competition, it is always very clear that the children and their coaches feel that they are competing. Some observers complain that the one problem they have with the OM competition is the "little league" spirit that is unfortunately very clearly observable in some of the parent coaches. The issue of competition may have a likely impact upon the effects of this study but were not addressed as an effect which could be measured in the design of this study.

In her case study of coaches and members of five OM teams, Cohen (1987) also found differences among students when asked to cite the advantages and disadvantages of working on a team. Some felt that sharing of work and responsibility was an advantage and had experienced support from their group. However, the most frequently cited disadvantages were:

> that it was difficult to compromise and reach consensus about ideas . . . they resented having their own ideas wiped out and found that it was difficult having to listen to one another . . . Some students . . . cited responsibility as a disadvantage . . . In OM more decision making power is given to the team members {by the OM coach than by athletic team coaches. The students in OM} like the feeling of group support but not the responsibility for others; they like the proliferation of ideas but not at the expense of their own ideas. Can these apparent paradoxical attitudes be ameliorated with education? It might be helpful if coaches and the OM organization itself trained the teams in decision making and group processes. Perhaps this training should go beyond OM team training. Decision making, group dynamics, and creative problem solving should be part of curriculum design for all students. It is essential in the design for programs for the gifted (Cohen, 1987, pp. 232-234).

Cohen's (1987) recommendation for further training in team process is appropriate to the population in the present study. Informal responses from children on the OM teams in this district are similar to those in Cohen's study. Evaluations of the parent coaches consistently indicated that the greatest needs were for increased coaches' training in discipline and keeping the team on target. Cohen's recommendations, then, are appropriately phrased in general terms for OM coaches. In the present study, the coaches received only two hours of training. It is likely the OM coaches would profit from further training in the team creative problem solving technique which focuses on acceptance of ideas and asking other team members to explain and elaborate on their ideas.

<u>Results Which Were Not Hypothesized</u>

<u>Changes in time--differences in pre- and posttest scores</u>. For the creative construct, a very strong effect of time was observed. The effect size of .75 of the multivariate effect indicates that a great amount of the variance in the multivariate creative construct is accounted for by differences in pre- and posttests. The univariate findings show strong decreases for verbal originality, N^2=.21, and a very weak decrease in verbal fluency, N^2=.01, and figural elaboration, N^2=.07. There is a strong increase in figural originality, N^2=.20, and a weak increase in Similes, .06, and in figural fluency, N^2=.04. The discussion of

these findings suggest that, in light of findings prevalent in research literature, the significant and unpredicted decreases in verbal TTCT creativity and in figural elaboration showed primarily to be attributed to the context effect and to regression toward the mean. In light of the increase in figural originality and in Similes and the context of the environment of the enrichment program which was interrupted for the lengthy testing of the study, the conclusion is reasonable that the verbal TTCT scores are a depressed effect specific to this study and that the result of lowered creativity is suspect. The increases in Similes, in figural originality, and in figural fluency could be attributed to overall effect of the enrichment program. In the setting of this study, it would not have been ethical or possible to obtain an equivalent control group of gifted children who were to be denied the service of the enrichment program. Therefore, no definitive answers can be derived, and these changes could simply have been caused by maturation.

Significant differences between pre- and posttests occurred also in the affective construct. The strength of association for the multivariate construct was only moderate, .15, and was a very weak but significant effect of increase in creative self-concept and in locus of control. These results of a significant but very weak beneficial effect on creative self-concept and locus of control may be attributed to the overall enrichment program or simply to maturation. As discussed above, it is rare to obtain results of change in affective measures associated with a creativity training program. It is likely that affective results, then, are of a weak effect size. Other conclusions related to affect and creativity are discussed above.

Because grade was discovered to be a significant predictor of several of the dependent creative variables, an analysis that had not been planned in the research design was used to test the effect of differences in grade on the creative construct. The analysis revealed that difference in grade, the interaction between grade levels over time, accounted for .76 of the variance in the creative construct. The strongest effects of differences between grades and pre- and posttest scores is noted for the univariate strength of association tests for closure, .18, and weaker effects for the other variables affected by this interaction. In the effect on verbal fluency, it was found that the very significant decrease of posttest scores was found primarily only in Grade 2+3 and verbal originality decreased for all grades, but less so for Grade 5. Likewise, the decline for posttest scores for figural elaboration was not evident in Grade 5. These results indicate that the context effect and/or regression toward the mean may have had the greatest effect at Grade 2+3, in particular with the younger children who had to leave the class to complete the individual administration of the verbal tests. The major reasons for decline in scores are discussed in relation to the context effect, the fourth grade slump, probable motivational effects of district-wide achievement testing for third grade students, and perhaps some effect of motivational differences between group and individual administration of the verbal TTCT for Grades 3 and younger.

Pre-existing differences between Odyssey of the Mind groups. The analyses found, based on pretest scores, that the OM groups were not equivalent on both the creative and affective constructs. The children who chose not to join an OM team tested lower in pretest cognitive creativity in Similes, closure, and on total of creative strengths than those who chose to be on a team. Moreover, those who were later found to be high effort had higher scores on Similes than either of the other two groups.

A similar pattern was evident for the affective construct and for creative self-concept. Creative self-concept pretest scores, similar to the cognitive creativity pretest scores for Similes and the creative strengths, were higher for the OM-hi group than either of the other groups. These results show that children who chose to be on a team held a higher self-concept with regard to their own creativity than children who chose not to join a team. Moreover, those who volunteered to participate in OM also were higher on their cognitive creativity scores. This finding is especially of interest to a major premise of this study because it parallels some of the results of Wright et al. (1975) in offering construct validation of Ideas as a measure of self-concept congruent with the tested creativity. These findings make sense, persons who consider themselves to be creative and whose measured creativity indicates that they are high in cognitive creative thinking abilities indeed would be and perhaps should be more likely to take the risk of entering a creativity competition than those whose measured creativity is lower.

Pretest differences for the prior experience variables. There were significant differences within the sample in tested creativity that were associated with prior experience. It is interesting to note that the same cluster of abilities appears associated for the children with prior experience in OM and for the children with more experience in the enrichment program. However, these variables are not independent because of considerable overlap, that is, often a child who has had the opportunity to have the experience of being on a team during the prior year or more by definition had to have been in the program for at least seven months.

Children who participated in the enrichment program for 13 months or more or who had at least one year of prior experience on an OM team showed higher scores on Similes, total of creative strengths, closure, figural originality, and elaboration, but lower scores on verbal flexibility. The lower scores on verbal flexibility are viewed as a chance occurrence because it makes little sense with the findings of the OM x Time interaction and with the presence of disproportionately high scores in the no-or-little experience and non-OM cell.

With the exception of the possibly chance effects of verbal flexibility, a cluster of creative thinking variables is associated with experience in the creativity training program of the enrichment program and/or with prior experience on an OM team. These findings lend support to the major hypothesis that participation on an OM team may enhance a child's creative thinking abilities.

Measurement and Creativity

Motivation and the context effect. The results affirm the need for users of creativity tests to be very sensitive to the impact of numerous possible variations in environment and testing conditions on children's performance on creativity tests. Overtesting by using too lengthy a battery was a problem which was greatly exacerbated by the effect of conducting the testing during classtime of favored activities. These effects are likely to account for a good portion of the decrease in verbal TTCT and figural elaboration scores which reflected the children's lowered motivation to expend their full creative energy on the test tasks.

Torrance Tests of Creative Thinking. In searching to determine the adequacy of inter-rater reliability for this study, a study within the study evolved which adds to the information base for reliability and validity of the TTCT streamlined scoring procedures as described above and in Appendix E. The results support the findings of other researchers who found it possible to obtain satisfactory inter-rater reliability correlations by careful study of the manual for the standard scoring of the TTCT. In this study, adequate inter-rater reliability was found with three other raters for the streamlined scoring of the figural tests. Raw scores must be used to compute inter-rater reliability values. Although the distributions were correlated, the scores were somewhat discrepant and most markedly different on elaboration, a finding consistent with the results of Halpin and Halpin (1974).

It was discovered that use of standard scores for the streamlined scoring introduced a source of unreliability to the closure and titles measures and possibly also to elaboration and the creative strengths. Therefore, even for researchers who are interested in examining relationships relevant to the different scales, it is likely that the raw scores will offer more reliable and, therefore, more valid measurement. If multivariate analysis is utilized, the raw scores can be used as readily as the standard scores. However, for users whose purpose is identification, the findings of this study strongly support the recommendation of Chase (1985) for a single score instead of use of the highly correlated verbal measures, and of Torrance and Ball (1984) for the Creativity Index.

Because the inter-rater reliabiity can vary on any open-ended test of divergent thinking, the use of a single rater and of local norms is strongly encouraged. Because of the variability attributable to the use of different raters and even more so because of the variability that is attributable to possible differences in amount of motivation the individual brings to the testing session, this author strongly concurs with Davis and Rimm (1985) that firm cut-off scores should not be used for purposes of identification and that scores on the TTCT should be used with other information. This author contends that, with the TTCT's adequate evidence of predictive validity, however, the TTCT is an adequate predictor if used with some stable measure of creativity such as a personality trait (see discussion in Appendix E, identification of the creatively talented).

<u>Creative self-concept: How Many Ideas?</u> The results of this study indicate that Ideas has adequate reliability and showed a fairly normal range of distribution in this sample of gifted children. It revealed changes in the children

which were consistent with the hypothesized effects of growth in creative self-concept over time and in relation to effort in an OM group. The results of intercorrelation and regression analyses support its base of validity. As discussed above under affective instruments, Ideas fills a gap and may be a promising instrument for investigation of self-concept specific to creativity.

Other measures. Similes and the Sears Inventory showed good reliability and yielded results that could be explained in relation to prior research and theoretical literature. The IAR, however, showed low reliability in use of the I+ and I- scales. Some of the items may have elicited ambiguous responses for reliable rating of internality of locus of control in this sample of gifted childred.

Suggestions for Further Research

Further analyses of the present data are possible and appropriate. Investigation of the strength of the relation found by the correlational data (see Tables P-25 and P-26, Appendix P) between creative and affective variables of this study is planned by use of canonical correlational analyses. Such analyses will further address the justifiability of examining creative and affective variables as separate constructs rather than in a single construct (Kolloff & Feldhusen, 1984).

The inequality of distribution of subjects in the different cells of this sample is large but not excessive (see Appendix B). An analysis in which subjects are randomly deleted from the non-OM condition in order to equalize the cells would simplify this complex design for multivariate analysis. If analyses with the cells equalized by random deletion of cases (Tabachnick & Fidell, 1983) confirms the results as analyzed in this dissertation, generalizability would be enhanced.

Further, exploration is appropriate to determine if results would be more clear if different measures available from these instruments were to be used. For the TTCT, abstractness of titles bore surprisingly little relationship to the verbal measures of the TTCT or to Similes. Moreover, titles was the only measure of the creativity battery for which no significant findings were obtained. It is possible that the lack of relationship may in large part be atributable to unreliability of the conversion to standard scores. In order to test the relationship of titles, but even more important, to determine the relationship, if any, of the figural measures to OM training, the data should be analyzed with use of raw scores.

For the affective instruments, some questions remain about whether subscale scores should have been used rather than the full scale scores. That is, these data would be a good source for exploration of the divergent thinking subscale on the Sears Inventory (see Chapter III). The finding of lowered scores for the OM-lo group in general self-concept might be further analyzed by examining the children's responses to those items on the Sears which address creative self-concept. Such analysis could, hopefully, shed light on whether the decrease was indeed a decrease in general self-concept or decrease in the divergent self-concept scale of the Sears. The IAR effort scale scores could be analyzed to determine if those scores obtain better reliability and possibly might bear a significant relationship to the OM effort group variables.

Some additional questions remain which might be explored with the present data. These relate to the type of OM problem selected by the team and the effects of competition. Teams worked on problems which were primarily either nonlanguage problems or problems which permitted free use of language. The training practice for brainstorming of the spontaneous problems would be primarily nonverbal or verbal for these two types of problems. In the year of the study, very few of the teams selected a nonverbal problem. An exploratory analysis could be performed to test for possible differences in verbal TTCT and figural TTCT patterns between teams that chose a language or nonlanguage OM problem.

It has been suggested as well that perhaps motivation bears an inverted U-shaped relationship to participation in creative tasks (K. S. Bull, personal communication, July, 1988, see Appendix A). That is, where participation or effort to engage in creative enterprise is expected to be minimal then the product(s) will be unlikely to be judged as creative. Peak creativity would be likely to occur under high levels of motivation and/or of effort. However, it is likely that excessive effort would result in too high a stress level and a freeze in the creative productivity. If so, it would be likely that the level for "excessive effort" would vary with the individual as well as with the task. It is, therefore, possible that for some of the children in the OM-hi effort group creative productivity could be lowered under extremely high levels of effort. Examination of the standard deviations of the posttest creativity measures prior to adjustment of scores (see Table P-2, Appendix P) indicate that such an interpretation might be plausible. That is, there is greater variability in the OM-hi group on posttest scores for Similes, verbal flexibility, figural fluency, and the total of creative strengths. There was also greater variability in the OM-hi group in posttest general self-concept (see Table 6). These results suggest further analysis to explore the interpretation of a possible inverted U-shaped relationship between effort and creative production.

Issues relating to the effects of competition cannot be adequately addressed by the design of this study. However, it is possible that some confounding might have occurred between high and low effort and winning versus losing. This confounding was, hopefully, minimized by obtaining the peer and self-ranking of effort the week prior to competition.

Some questions, however, still lurk: How many of those who were high effort were actually on teams that won? Is there a relation between low and high effort as measured in this study and winning at the competition? Did the effort scale categorize children totally across teams, or were there some teams, because of their weighting of the amount of the team's work, in which most members were rated as low effort?

Future studies are necessary using a different sample of subjects to determine whether the results of this exploratory study may be replicated. There is a need for future research to conduct a study which controls the numerous uncontrolled variables found in this quasi-experimental study. Rather than using an intact sample as in the present study, random sampling or selection techniques should be used. A true experimental paradigm could be used to test the effect for students who are in OM and those who are not. It would be helpful to determine if the results would be different in a study of OM where the coaches are teachers or have more structured training in facilitating the OM CPS process than was available to the coaches in this study.

The findings of use of effort as an independent variable suggest many fruitful areas of future research. Consideration of differences in amount of effort exerted in the creative enterprise of OM was shown to account for differences in measured creativity and in affect in this study. It is likely that differences in individual effort may be a helpful variable to clarify previously conflicting results about the effectiveness of creativity training. A longitudinal study could be designed to assess the creative and affective changes in children who participate on OM teams. If the measurement was conducted only once a year it would be likely to minimize the detrimental effects of testing in the context of an enrichment program.

Implications of the Findings

Implications of the findings with respect to use of the TTCT are presented in the section on measurement. The results of the methods used in this study suggest that researchers should use appropriate specific measures. The new scale of creative self-concept (Ideas) showed promise as a reliable and sensitive instrument that was sensitive to the hypothesized findings of this study. It is possible that changes occurred in this study on a general measure of self-concept because the Sears contained items specific to creativity. It is possible that the Sears might be more sensitive than the Piers-Harris (Piers, 1977) as a useful general self-concept instrument in studies of gifted elementary children where the curriculum has a considerable component on creativity. Further investigation is again warranted to answer such questions.

A very practical recommendation from this study would be that coaches/ facilitators of team problem solving situations need to pay particular attention to helping to facilitate effort and/or participation in children who may tend to hold back in the group. Team members have been observed to learn the process of withholding judgment very quickly. For example, coaches could teach their teams to delay judgment during the brainstorming phase by instructing students in the following manner: "If you do not understand or like your friend's ideas, you are not allowed to judge it during brainstorming. You can, instead, help your friend by saying, 'Please explain your idea more completely.'" When team members become skilled in the CPS techniques of deferring judgment, the team's environment is appropriate for enhancing creativity (Parnes, 1972).

Importance of the Study

This study offers five new contributions to creativity research.
1) Validation is provided for the new figural streamlined scoring procedures of the TTCT (Torrance & Ball, 1984) by multivariate comparison of each of the figural scores with each other, with the verbal TTCT measures, and with another verbal creativity measure. The scoring categories of closure and total of creative strengths were shown to bear a relationship with prior experience with training in creativity as was the figural Creativity Index. 2) The study shows that

multivariate analysis of the various components of creativity and of affect is helpful to show strength of relation between the various components of these two constructs. Multivariate analysis is the appropriate method to control Type I error rate (Larrabee, 1982), and is helpful to examine the relationships among creative variables (Treffinger & Poggio, 1972).

3) Changes in affect were assessed by use of a specific measure of locus of control and an appropriate newly developed scale of creative self-concept rather than by a general self-concept measure with content unrelated to creativity. Changes of affect were obtained in those two measures as well as in the general self-concept scale with content relevant to behaviors emphasized in the gifted program. The brief scale of creative self-concept, developed for this study, was shown to reflect positive increase over time in children in the program. These findings support the work of Shavelson and Bolus (1982) and extend their contributions to the domain of affect specific to creativity. Fostering of creativity and self-directed learning were among the important goals of the gifted program, and it was found that creative self-concept and internality of locus of control were higher at the end of the year than at the beginning. In this study, the hypothesized changes in affect associated with the OM CPS training program were found by use of appropriate, specific measures of affect. These findings represent a new and important contribution to the literature on creativity training and affect. The findings suggest the need to use measures appropriate and specific to the content of affect associated with creative and gifted programs in future evaluation and research studies.

4) A method of defining effort as an independent variable was devised which successfully separated subjects in the study into meaningful groups. The research design, which accounted for differences in the amount of effort children invested in their creative work clarified the relationship of participation and

nonparticipation with the dimension of task commitment or effort involved in the participation. 5) Finally, empirical evidence was provided to test the hypothesis that team CPS, in particular OM as a form of team CPS, is effective in increasing the creativity and affect in gifted children. It was found that the children who were high in effort in their OM work showed higher creativity and creative self-concept than children who were not on a team and those who were on a team but exerted little effort in their team's work.

Summary

In this chapter the results were discussed for the hypotheses of the study. The findings of pre-existing differences between the OM groups were interpreted. The effects of the OM treatment, of pre-existing differences between groups, and of changes in time across groups were discussed in relation to light they might shed on the validity of the various dependent variables. Several issues bearing on these findings and their relation to research in pertinent areas of creativity were discussed. Issues of reliability, validity, selection, and use of the various instruments of the study were discussed in relation to the findings of this study. The relation of the results to creativity research was discussed for the general issues of effort, social influences, affect, motivation, intelligence, and grade level. The new measure of creative self-concept was discussed. Suggestions for future research and implications of the study were offered.

In summary, five new contributions to creativity research were included in this study: 1) Further validation of the figural streamlined scoring of the TTCT; 2) Confirmation of the usefulness of multivariate statistical procedures for research in creativity and affect; 3) Development and validation of a new creative self-concept scale appropriate for elementary gifted children; 4) Identification and use of effort as an independent variable; and 5) Evidence that team CPS in the form of OM is beneficial for gifted children, for those who are high in their effort in their work on OM teams.

REFERENCES

Abramson, J. H. (1976). The effects of non-competitive and group competitive situations on the verbal and figural and creativity of college students. Dissertation Abstracts International, 36, 83A.

Amabile, T. M. (1983). The social psychology of creativity. New York: Springer-Verlag.

Arieti, S. (1976) Creativity: The magic synthesis. New York: Basic Books.

Artley, N. L., Van Horn, R., Friedrich, D. D., & Carroll, J. L. (1980). The relationship between problem finding, creativity and cognitive style. The Creative Child and Adult Quarterly, 5, 20-26.

Ball, O. E. & Torrance, E. P. (1984). Torrance Tests of Creative Thinking: Streamlined scoring workbook Figural A. Bensenville, IL: Scholastic Testing Service.

Barron, F. (1963). Creativity and psychological health: origins of personal vitality and creative freedom. Princeton: Van Nostrand.

Beck, P. J. (1979). Locus of control and task instruction effects upon creative problem solving. Dissertation Abstracts International, 40, 1351A.

Bennett, L. E. (1984). Achievement differences in gifted children: Possible non-intellective factors. Dissertation Abstracts International, 45, 487A (University Microfilms No. 84-11844).

Bennett, O. G. (1982). An investigation into the effects of creative experience in drama upon the creativity self-concept and achievement of fifth and sixth grade students. Dissertation Abstracts International, 43, 1809-A.

Besemer, S., & O'Quin, K. (1986). Analyzing creative products: Refinement and test of a judging instrument. Journal of Creative Behavior, 20, 115-126.

Besemer, S. P., & Treffinger, D. J. (1981). Analysis of creative products: Review and synthesis. Journal of Creative Behavior, 15, 158-178.

Blumenfield, P. C., Pintrich, P. R., & Hamilton, V. L. (1986). Children's concepts of ability, effort, and conduct. American Educational Research Journal, 23, 95-104.

Bolen, L. M., & Torrance, E. P. (1978). The influence on creative thinking of locus of control, cooperation, and sex. Journal of Clinical Psychology, 34, 903-907.

Bradley, R. H., & Gaa, J. P. (1977). Domain specific aspects of locus of control: Implications for modifying locus of control orientation. Journal of School Psychology, 15, 18-24.

Bray, J. H. & Maxwell, S. E. (1985). Multivariate analysis of variance. Beverly Hills, CA: Sage.

Bull, K. S. (1978). Maximizing creative potential. In J. Ferver, R. E. Clasen & B. Robinson (Eds.), Simple gifts: The education of the gifted, talented and creative (pp. 211-228). (2nd ed.). Madison, WI: UWEX.

Bull, K. S., & Fishkin, A. S. (1987). OK-OM coaches training manual. (rev. ed.) (Available from OK-OM, E. Hobbs, Rt. 7, P. O. Box 594, Ada, OK 74820.)

Burk, E. A. (198). Relationship of temperamental traits to achievement and adjustment in gifted children. Dissertation Abstracts International, 40, 5784A.

Burns, D. E. (1988). The effects of group training activities on students' creative productivity. Dissertation Abstracts International, 48, 3072A.

Callahan, C. M. (1978). Developing creativity in the gifted and talented: What research and experience say to the teacher of exceptional children. Reston, VA: Council for Exceptional Children.

Carmines, E. G. & Zeller, R. A. (1979). Reliability and validity assessment. Beverly Hills, CA: Sage.

Carroll, J. L., & Laming, L. R. (1974). Giftedness and creativity: Recent attempts at definition: A literature review. Gifted Child Quarterly, 18, 85-96.

Casler, J. M. (1983). Divergent production measures and self-concept among third, fourth, and fifth grade students in open and traditional classrooms. (Doctoral dissertation, Oklahoma State University, Stillwater, 1982.) Dissertation Abstracts International, 43, 3265A.

Chase, C. I. (1985). Review of the Torrance Tests of Creative Thinking. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook: Vol. II (pp. 1631-1632). Lincoln, NE: University of Nebraska-Lincoln, Buros Institute of Mental Measurements.

Christie, T. (1976). Environmental factors in creativity. In A. M. Biondi and S. J. Parnes Assessing creative growth (pp. 185-205). Buffalo, NY: Bearly Limited.

Clark, B. (1983). Growing up gifted: Developing the potential of children at home and at school (2nd ed.). Columbus: Charles E. Merrill.

Clark, P. M. (1978). Review of Similes Test. In O. K. Buros (Ed.), The eighth mental measurements yearbook: Vol. I (pp. 371-372). Highland Park, NJ: Gryphon Press.

Cohen, G. F. (1987). Creatively gifted students: A case study of teams participating in the Olympics of the Mind competition. Dissertation Abstracts International, 47, 2867A. (University Microfilms No. 86-20344).

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev. ed.). New York: Academic Press.

Cohen, J. & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohn, C. M. G. (1985). Creativity training effectiveness: A research synthesis. Dissertation Abstracts International, 45, 2501A.

Coleman, J. M., & Fults, B. A. (1982). Self-concept and the gifted classroom: The role of social comparisons. Gifted Child Quarterly, 26, 116-120.

Companik, P. J. (1978). Assessment of the attempts to modify locus of control orientation in preadolescent school children. Dissertation Abstracts International, 38, 3386-3387B.

Coopersmith, S. (1967). The antecedents of self-esteem. Palo Alto, CA: Consulting Psychologists Press.

Cottrell, N. B. (1968). Performance in the presence of other human beings: Mere presence, audience, and affiliation effects. In E. C. Simmel, R. A. Hoppe & G. A. Milton (Eds.), Social facilitation and imitative behavior (pp. 91-111). Boston: Allyn & Bacon.

Crabbe, A. B. (1982). Creating a brighter future: An update on the Future Problem Solving Program. Journal for the Education of the Gifted, 5, 2-11.

Crabbe, A. B. (1985). The coach's guide to the Future Problem Solving Program. Laurinburg, NC: The Future Problem Solving Program.

Crandall, V. C. (1978, August). New developments with the Intellectual Achievement Responsibility Scale. In The Goal-specific locus of control scales. Symposium conducted at the 86th Annual Convention of the American Psychological Association, Toronto, Canada.

Crandall, V. C., Katkovsky, W., & Crandall, V. J. (1965). Children's belief in their own control of reinforcements in intellectual-academic achievement situations. Child Development, 36, 91-109.

Crandall, V. J., Katkovsky, W., & Preston, A. (1962). Motivational and ability determinants of young children's intellectual achievement behaviors. Child Development, 33, 643-661.

Crittenden, M. R., Kaplan, M. H., & Heim, J. K. (1984). Developing effective study skills and self-confidence in academically able young adolescents. Gifted Child Quarterly, 28, 25-30.

Cronbach, L. J. (1960). Essentials of psychological testing. (2nd ed.). New York: Harper & Brothers.

Csikszentmihalyi, M., & Getzels, J. W. (1971). Discovery oriented behavior and originality of creative products: A study with artists. Journal of Personality and Social Psychology, 19, 47-52.

Davidman, L. (1982). Expressive encounters and the unleashing of creative potential. Gifted Child Quarterly, 26, 57-62.

Davis, G. A. (1983). Creativity is forever. Dubuque, IA: Kendall/Hunt.

Davis, G. A. & Rimm, S. B. (1985). Education of the gifted and talented. Englewood Cliffs, NJ: Prentice-Hall.

Davis, G. A., & Scott, J. A. (1971). Training creative thinking. New York: Holt, Rinehart and Winston.

Davis, G. A., & Subkoviak, M. J. (1975). Multidimensional analysis of a personality-based test of creative potential. Journal of Educational Measurement, 12, 37-43.

Delisle, J. R. (1985). Quasi-experiments in gifted child education: A confounding problem. Journal for the Education of the Gifted, 7, 151-155.

Dunn, R. S., & Price, G. E. (1980). The learning style characteristics of gifted students. Gifted Child Quarterly, 24, 33-36.

Eberle, B., & Stanish, B. (1980). CPS for kids: A resource book for teaching creative problem solving to children. Buffalo: D.O.K.

Elkind, D., Deblinger, J., & Adler, D. (1970). Motivation and creativity: The context effect. American Educational Research Journal, 7, 351-357.

Elkind, S. N. (1969). The relationship of orientation toward achievement to convergent and divergent problem-solving. Dissertation Abstracts International, 30, 1356-1357B.

Feldhusen, J. F., & Clinkenbeard, P. R. (1986). Creativity instructional materials: A review of research. Journal of Creative Behavior, 20, 153-182.

Feldhusen, J. F., Speedie, S. M. & Treffinger, D.J. (1971). The Purdue Creative Thinking Program: Research and evaluation. NSPI Journal, 10, (3), 5-9.

Feldhusen, J. F., & Treffinger, D. J. (1985). Creative thinking and problem solving in gifted education (3rd ed.). Dubuque, IA: Kendall/Hunt.

Finn, J. D., & Mattsson, I. (1978). Multivariate analysis in educational research: Applications of the multivariate program. Chicago: National Educational Resources.

Firestien, R. L., & Treffinger, D. J. (1983). Creative problem solving: Guidelines and resources for effective facilitation. G/C/T, 26, 2-10.

Fishkin, A. S. (1987a, March). Affective changes and OM team creative problem solving in the gifted. Paper presented at the annual spring conference of the Gifted Students Institute, Arlington, TX.

Fishkin, A. S. (1987b, November). Assessing creative and affective changes in gifted children. Paper presented at the annual convention of the National Association for Gifted Children, New Orleans.

Foster, G. W. (1981). Creativity and the group problem solving process. Dissertation Abstracts International, 42, 3093A.

Fowler, J. W., & Peterson, P. L. (1981). Increasing reading persistence and altering attributional style of learned helpless children. Journal of Educational Psychology. 73, 251-260.

Franks, B., & Dolan, L. (1982). Affective characteristics of gifted children: Educational implications. Gifted Child Quarterly, 26, 172-177.

Frederiksen, N. (1978). Review of Similes Test. In O. K. Buros (Ed.), The eighth mental measurement yearbook: Vol. I (pp. 372-373). Highland Park, NJ: Gryphon Press.

Freud, S. (1908). Creative writers and day-dreaming. In P. E. Vernon (Ed.), Creativity: Selected readings. (pp. 126-136). Harmondsworth, England: Penguin.

Friedman, P. G., Friedman, R. J., & Van Dyke, M. (1984). Identifying the leadership gifted: Self, peer, or teacher nominations. Roeper Review, 7, 91-94.

Fults, E. A. (1981). The effect of an instructional program on the creative thinking skills, self-concept and leadership of intellectually and academically gifted elementary students. Dissertation Abstracts International, 41, 2931A.

Gabrielle, R. V. (1985). Identification as gifted and talented: Effects on internal-external control of intellectual achievement and self-concept of ability. Dissertation Abstracts International, 46, 896A.

Gallagher, J. J. (1975). Teaching the gifted child (2nd ed.). Boston: Allyn & Bacon.

Getzels, J. W. (1975). Problem finding and the inventiveness of solutions. Journal of Creative Behavior, 9, 12-18.

Getzels, J. W., & Jackson, P. W. (1962). Creativity and intelligence: Explorations with gifted students. New York: Wiley.

Goff, K. (1987). Definition of creativity: OM versus non-OM. Unpublished master's thesis, Oklahoma State University, Stillwater.

Gordon, D. A., Jones, R. H., & Short, N. L. (1977). Task persistence and locus of control in elementary school children. Child Development, 48, 1716-1719.

Gourley, T. J., & Micklus, C. S. (1978). Olympics of the Mind: Coaches' training manual. Glassboro, NJ: Creative Competitions.

Gowan, J. C. (1972). The development of the creative individual. San Diego: Knapp.

Gowan, J. C. (1977). Some new thoughts on the development of creativity. Journal of Creative Behavior, 11, 77-90.

Gray, J. S. (1986). Creative drama with four-grade students: Effects of using guided imagery. Dissertation Abstracts International, 47, 772A.

Guidry, L. J. & Knight, D. F. (1976). Comparative readibility four formulas and Newbery books. Journal of Reading, 19, 522-556.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Halpin, G. & Halpin, G. (1974). Can self trained scorers reliably score the Torrance Tests of Creative Thinking? Psychology in the Schools, 11, 56-58.

Hare, A. P. (1981). Creativity in small groups. Beverly Hills, CA: Sage.

Harrington, D. M., Block, J., & Block, J. H. (1983). Predicting creativity in preadolescence from divergent thinking in early childhood. Journal of Personal and Social Psychology, 45, 609-623.

Harrington, P. (1984). An examination of the creative problem-solving model Olympics of the Mind and its effects on the attitudes of teachers and students as well as its effects on the problem-solving abilities of gifted students. Unpublished master's thesis, Ohio State University, Columbus.

Harris, R. J. (1975). A primer of multivariate statistics. New York: Academic Press.

Heist, P. (1968). Creative students: College transients. In P. Heist (Ed.), The creative college student: An unmet challenge (pp. 35-55). San Francisco: Josey-Bass.

Hendrickson, P. (1986). Effects of three creative problem solving treatments on products of native and other gifted Alaskan students. (Doctoral dissertation, University of Washington, Seattle, 1985). Dissertation Abstracts International, 46, 2904A-2905A.

Hocevar, D. (1981). Measurement of creativity: Review and critique. Journal of Personality Assessment, 45, 450-463.

Holen, M. C., & Newhouse, R. C. (1976). Large versus small group administration of selected attitude, self-concept, and locus-of-control scales. Measurement and evaluation in guidance, 8, 255-257.

Hoomes, E. W. (1984). Future Problem Solving: Preparing students for today and tomorrow. G/C/T, 32, 15-18.

Houtz, J. C., Montgomery, C., Kirkpatrick, L., & Feldhusen, J. F. (1979). Relationship among measures of evaluation ability (problem solving), creative thinking and intelligence. Contemporary Educational Psychology, 4, 47-54.

Houtz, J. C., Rosenfield, S., & Tetenbaum, T. J. (1978). Creative thinking in gifted elementary children. Gifted Child Quarterly, 22, 513-519.

Houtz, J. C., & Speedie, S. M. (1978). Processes underlying divergent thinking and problem solving. Journal of Educational Psychology, 70, 848-854.

Howieson, N. (1981). A longitudinal study of creativity: 1965-75. Journal of Creative Behavior, 15, 117-134.

Isaac, S., & Michael, W. B. (1981). Handbook in research and evaluation (2nd ed.). San Diego: EdITS.

Isaksen, S. G., Stein, M. I., Hills, D. A., & Gryskiewicz, S. S. (1984). A proposed model for the formulation of creativity research. Journal of Creative Behavior, 18, 67-75.

Isaksen, S. G. & Treffinger, D. J. (1984). Creative problem solving: The basic course. Buffalo, NY: Bearly Limited.

Janos, P. M., Fung, H. C., & Robinson, N. M. (1985). Self-concept, self-esteem, and peer relations among gifted children who feel "different." Gifted Child Quarterly, 29, 78-82.

Jones, W. C. (1983). Effects of time spent in an enrichment program for gifted and talented students (Doctoral Dissertation, Pennsylvania State University, 1982). Dissertation Abstracts International, 43, 2219A.

Kanoy, R. C., Johnson, B. W., & Kanoy, K. W. (1980). Locus of control and self-concept in achieving and under-achieving bright elementary students. Psychology in the Schools, 17, 395-399.

Khatena, J. (1982). Myth: Creativity is too difficult to measure! Gifted Child Quarterly, 26, 21-23.

Koberg, D., & Bagnall, J. (1976). The universal traveler: A soft-systems guide to: Creativity, problem-solving and the process or reaching goals. Los Altos: William Kaufmann.

Kogan, N., & Pankove, E. (1974). Long-term predictive validity of divergent-thinking tests: Some negative evidence. Journal of Educational Psychology, 66, 802-810.

Kolloff, P., & Feldhusen, J. F. (1984). The effects of enrichment on self-concept and creative thinking. Gifted Child Quarterly, 28, 53-57.

Krathwohl, D., Bloom, B., & Masia, B. (1964). Taxonomy of educational objectives. Handbook II: Affective domain. New York: David McKay.

Larrabee, M. J. (1982). Reexamination of a plea for multivariate analyses. Journal of Counseling Psychology, 29, 180-188.

Leary, M. R. & Altmaier, E. M. (1980). Type I error in counseling research: A plea for multivariate analyses. Journal of Counseling Psychology, 27, 611-615.

Lefcourt, H. M. (1966). Internal versus external control of reinforcement. Psychological Bulletin, 65, 206-220.

Linton, M. & Gallo, P. S. (1975). The practical statistician: Simplified handbook of statistics. Belmont, CA: Wadsworth.

MacKinnon, D. W. (1968). A dialogue on creativity. In P. Heist (Ed.), The creative college student: An unmet challenge (pp. 3-17.). San Francisco: Josey-Bass.

MacKinnon, D. W. (1975). IPAR's contribution to the conceptualization and study of creativity. In I. A. Taylor & J. W. Getzels (Eds.), Perspectives in creativity (pp. 60-89). Chicago: Aldine.

Maddox, C. D., Scheiber, L. M., & Bass, J. E. (1982). Self-concept and social distance in gifted children. Gifted Children Quarterly, 26, 77-81.

Maker, C. J. (1982). Teaching models in education of the gifted. Rockville, MD: Aspen Press.

Mansfield, R. S., Busse, F. B., & Krepelka, E. J. (1978). The effectiveness of creativity training. Review of Educational Research, 48, 517-536.

Marx, R. W. & Winne, P. H. (1978). Construct interpretations of three self-concept inventories. American Education Research Journal, 15, 99-109.

Maslow, A. (1962). Toward a psychology of being. Princeton: Van Nostrand.

Matlin, M. W., & Zajonc, R. B. (1968). Social facilitation of word associations. Journal of Personality and Social Psychology, 10, 455-460.

McCall, R. B., & Appelbaum, M. I. (1973). Bias in the analysis of repeated-measures designs: Some alternative approaches. Child Development, 44, 401-415.

McKenna, H. K. H. (1981). Relationship between perception of responsibility of academic achievement and actual achievement and effort in gifted, bright and average students. (Doctoral dissertation, University of Southern California, 1981). Dissertation Abstracts International, 42, 2377A.

McQuilkin, C. E. (1981). A comparison of personal and social concepts of gifted elementary students in different school programs. (Doctoral Dissertation, Kent State University, 1980). Dissertation Abstracts International, 41, 3530A.

Meadow, A., & Parnes, S. J. (1959). Evaluation of training in creative problem solving. Journal of Applied Psychology, 43, 189-194.

Meeker, M. N. (1969). The structure of intellect: Its interpretation and uses. Columbus, OH: Charles E. Merrill.

Micklus, C. S., & Micklus, C. (1987). Odyssey of the Mind program handbook: Instructional manual for teams and coaches. Glassboro, NJ: Odyssey of the Mind.

Milgram, R. M., & Milgram, N. A. (1976). Self-concept as a function of intelligence and creativity in gifted Israeli children. Psychology in the Schools, 13, 91-96.

Miller, M. B. (1983). An evaluation of an elementary gifted program. Dissertation Abstracts International, 43, 3475A. (University Microfilms No. 8307957)

Mourad, S. A. (1976). A factor analytic study of the streamlined scoring of Figural Form A of the Torrance Tests of Creative Thinking. Masters Abstracts, 1976, University of Georgia.

Moyers, B. (Producer). (1981). Creativity with Bill Moyers. (3rd in the series). (Videotape). New York: CBS-TV.

Myerson, E. S. (1981). Academic, affective, and aesthetic outcomes of a sixth grade creative dramatics program: a quantitative and qualitative evaluation. Dissertation Abstracts International, 42, 1445-A.

Nunnally, J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill.

Osborn, A. F. (1957). Applied imagination: Principles and procedures of creative thinking (rev. ed.). New York: Charles Scribner's.

Osborne, W. L., & LeGette, H. R. (1982). Sex, race, grade level and social class differences in self concept. Measurement and evaluation in guidance, 14, 195-201.

Parker, J. P. (1983). The leadership training model. G/C/T, 29, 8-13.

Parnes, S. J. (1961). Effects of extended effort in creative problem-solving. Journal of Educational Psychology, 52, 117-122.

Parnes, S. J. (1972). Creative potential and the educational experience. In S. J. Parnes, Creativity: Unlocking human potential. (pp. 6-30). Buffalo, NY: D.O.K. (Reprinted from paper presented at the annual convention of the American Psychological Association, Washington, D.C., April, 1967.)

Parnes, S. J. (1975). Aha! In I. A. Taylor & J. W. Getzels (Eds.), Perspectives in creativity (pp. 224-248). Chicago: Aldine.

Parnes, S. J., & Brunelle, E. A. (1967). The literature of creativity (part 1). Journal of Creative Behavior, 1, 52-109.

Parnes, S. J., & Meadow, A. (1960). Evaluation of persistence of effects produced by a creative problem-solving course. Psychological Reports, 7, 357-361.

Parnes, S. J., Noller, R. B., & Biondi, A. M. (1977). Guide to creative action: Revised edition of creative behavior guidebook. New York: Charles Scribner's.

Pearson, P. H. & Maddi, S. R. (1966). The Similes Preference Inventory: Development of a structural measure of the tendency toward variety. Journal of Consulting Psychology, 30, 301-308.

Piers, E. V. (1977). Piers-Harris Self-Concept Scale: Research monograph #1. Nashville, TN: Counselor Recordings and Tapes.

Purifico, R., & Micklus, C. (1987). OM judges' rulebook. Glassboro, NJ: Creative Competitions.

Reese, H. W., & Parnes, S. J. (1970). Programming creative behavior. Child Development, 41, 413-423.

Reese, H., Parnes, S., Treffinger, D., & Kaltsounis, G. (1976). Effects of a creative studies program on Structure of Intellect factors. Journal of Educational Psychology, 68, 401-410.

Renzulli, J. (1973). New directions in creativity. Mark II. New York: Harper & Row.

Renzulli, J. S. (1977). The Enrichment Trial Model. Mansfield, CT: Creative Learning Press.

Renzulli, J. S. (1978). What makes giftedness? Reexamining a definition. Phi Delta Kappan, 60, 180-184, 261.

Riley, J., & Parr, L. (1988). A cooperative learning model for creative problem solving. Challenge: Reaching and teaching the gifted child, 6, 41-42.

Rimm, S. B. (1980). The group inventory for finding creative talent (GIFT). Watertown, WI: Educational Assessment Service.

Rimm, S. B. (1986). Underachievement syndrome: Causes and cures. Watertown, WI: Apple Publishing Co.

Roe, A. (1952). A psychologist examines sixty-four eminent scientists. In P. E. Vernon (Ed.), Creativity: selected readings (pp. 43-51). Harmondsworth, England: Penguin.

Rose, L. H., & Lin, H. (1984). A meta-analysis of long-term creativity training programs. Journal of Creative Behavior, 18, 11-22.

Rosenthal, A., DeMers, S. T., Stillwell, W., Graybeal, S., & Zins, J. (1983). Comparison of inter-rater reliability on the Torrance Tests of Creative Thinking for gifted and nongifted students. Psychology in the Schools, 20, 35-40.

Ross, A., & Parker, M. (1980). Academic and social self concepts of the academically gifted. Exceptional Children, 47, 6-10.

Rotter, J. B. (1982). Generalized expectancies for internal versus external control of reinforcement. In J. B. Rotter (Ed.), The development and applications of social learning theory: Selected papers. (pp. 171-216). New York: Praeger. (Reprinted from Psychological Monographs, 80 (1, 1-28).)

Rotter, J. B. (1982). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. In J. B. Rotter (Ed.), The development and applications of social learning theory: Selected papers. (pp. 265-283). New York: Praeger. (Reprinted from Journal of Consulting and Clinical Psychology, 43, 56-57.)

Rotter, J. B. (1982). Introduction. In J. B. Rotter (Ed.), The development and applications of social learning theory: Selected papers. (pp. 1-12). New York: Praeger.

Roweton, W. E. (1982). Creativity and competition. Journal of Creative Behavior, 16, 89-96.

Rush, R. T. (1985). Assessing readibility: Formulas and alternatives, Reading Teacher, 39, 274-83.

Sasfy, J. H. (1976, April). Some relationships between causality for task persistence and the causal attribution of task outcomes. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Schack, G. D. (1987). Creative productivity and self-efficacy in children. Dissertation Abstracts International, 47, 3375A-3376A.

Schaefer, C. E. (1971). Similes. Port Huron, MI: Research Psychologists Press.

Schubert, D. S. P., & Biondi, A. M. (1977). Creativity and mental health: Part III - Creativity and adjustment. Journal of Creative Behavior, 11, 186-197.

Seagoe, M. (1975). Some learning characteristics of gifted children. In R. Martinson (Ed.), The identification of the gifted and talented. (pp. 20-21). Reston, VA: The Council for Exceptional Children.

Sears, P. S. (1963). The effect of classroom conditions on the strength of achievement motive and work output of elementary school children. V. 5. Department of HEW, Office of Education, Cooperative Research Project No. 873. Stanford, CA: Stanford University.

Sears, P. S. (1972). Effective reinforcement for achievement behaviors in disadvantaged children: The first year (Report No. TR-30). Washington, DC: Office of Education.

Sears, P. S. (1975). Memorandum with respect to use of the Sears Self-Concept Inventory. Princeton: Educational Testing Service Collection.

Sears, P. S., & Sherman, V. S. (1964). In pursuit of self-esteem. Belmont: Wadsworth Publishing Co.

Shaffer-Zamari, M. (1983). Effects of an instructional program and teacher's verbal praise on the creative performance of fourth grade children. Dissertation Abstracts International, 44, 1689A.

Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. Journal of Educational Psychology, 74, 3-17.

Shavelson, R. J., Hubner, J. J., & Stanton, J. C. (1976). Self-concept: Validity of construct interpretations. Review of Educational Research, 46, 407-441.

Sherman, M. S. B. (1977). Selected affective characteristics and creative problem-solving performance in gifted elementary-school children (Doctoral Dissertation, Fordham University, 1977). Dissertation Abstracts International, 38, 186A.

Shivley, J. E., Feldhusen, J. F., & Treffinger, D. J. (1972). Developing creativity and related attitudes. The Journal of Experimental Education, 41, (2), 63-69.

Slavin, R. E. (1986). Student team learning: An overview and practical guide (rev. ed.). Washington, D.C.: National Education Association.

Slavin, R. E., & Karweit, N. L. (1981). Cognitive and affective outcomes of an intensive student team learning experience. Journal of Experimental Education, 50, 29-35.

Speedie, S. M., Treffinger, D. J., & Houtz, J. C. (1976). Classification and evaluation of problem solving tasks. Contemporary Educational Psychology, 1, 52-75.

SPSS, Inc. (1986). SPSS^{x} user's guide. (2nd ed.). Chicago: Author.

Starko, A. J. (1988). Effects of the revolving door identification model on creative productivity and self-efficacy. Gifted Children Quarterly, 32, 291-297.

Stein, M. I. (1968). Creativity. In E. F. Borgatta & W. W. Lambert (Eds.), Handbook of personality theory and research (pp. 900-942). Chicago: Rand McNally.

Stein, M. I. (1974-1975). Stimulating creativity (Vols. 1-2). New York: Academic Press.

Sund, R. B., & Carin, A. (1978). Creative questioning and sensitive listening techniques: A self-concept approach (2nd ed.). Columbus, OH: Charles E. Merrill.

Tabachnick, B. G., & Fidell, L. S. (1983). Using multivariate statistics. New York: Harper and Row.

Tadlock, S. E. (1981). The effect of a basic movement education program in the creative thinking skill and self-concept of gifted students. Dissertation Abstracts International, 41, 3407A.

Tallent, M. K. (1985). Effects of the Future Problem Solving Program on gifted students' ability to solve futuristic problems. (Doctoral dissertation, Texas A & M University, College Station, 1985). Dissertation Abstracts International, 47, 843A.

Tallmadge, G. F. (1988). Book review {Review of Advances in quasi-experimental design and analyses.} Educational Evaluation and Policy Analysis, 10, 71-73.

Tannenbaum, A. J. (1983). Gifted children: Psychological and educational perspectives. New York: Macmillan.

Taylor, I. A. (1975). A retrospective view of creativity investigation. In I. A. Taylor & J. W. Getzels (Eds.), Perspectives in Creativity (pp. 1-36). Chicago: Aldine.

Tetenbaum, T., & Houtz, J. (1978). The role of affective traits in the creative and problem-solving performance of gifted urban children. Psychology in the Schools, 15, 27-32.

Title 70, Oklahoma Statutes, 1987, Sections 1210.301; et seq.

Torrance, E. P. (1972a). Can we teach children to think creatively? Journal of Creative Behavior, 6, 114-143.

Torrance, E. P. (1972b). Predictive validity of the Torrance Tests of Creative Thinking. Journal of Creative Behavior, 6, 236-252.

Torrance, E. P. (1974a). Torrance Tests of Creative Thinking: Directions manual and scoring guide. (Verbal Booklet A-Verbal Booklet B). Bensonville, IL: Scholastic Testing Service.

Torrance, E. P. (1974b). Torrance Tests of Creative Thinking: Norms-technical manual. Bensenville, IL: Scholastic Testing Service.

Torrance, E. P. (1978). Giftedness in solving future problems. Journal of Creative Behavior, 12, 75-86.

Torrance, E. P. (1979). The search for satori and creativity. Buffalo, NY: Creative Education Foundation.

Torrance, E. P. (1984). Selection from Healing qualities of creative behavior. In J. P. Torrance (Ed.), Over the years: Research insights of E. Paul Torrance. (p. 16). Athens, GA: University of Georgia, Torrance Center for Gifted, Creative and Future Behaviors. (Reprinted from Gifted Child and Adult Quarterly, 1978, 3, 146-158.)

Torrance, E. P. (1984). Selection from Some implications of creativity research for gifted education. In J. P. Torrance (Ed.), Over the years: Research insights of E. Paul Torrance. (p. 20). Athens, GA: University of Georgia, Torrance Center for Gifted, Creative and Future Behaviors. (Reprinted from Creativity Research and Educational Planning, 1982, Ventura, CA: Ventura County Superintendent of Schools.)

Torrance, E. P. (1985). Forward. In A. B. Crabbe, The coach's guide to the Future Problem Solving Programs. (pp. 1-3). Laurinburg, NC: The Future Problem Solving Program.

Torrance, E. P. (1987). Guidelines of administration and scoring/comments on using the Torrance Tests of Creative Thinking. Bensenville, IL: Scholastic Testing Service.

Torrance, E. P., & Ball, O. E. (1984). Torrance Tests of Creative Thinking: Streamlined (revised) manual including norms and directions for administering and scoring Figural A and B. Bensenville, IL: Scholastic Testing Service.

Torrance, E. P., & Ball, O. E. (in press). Torrance Tests of Creative Thinking: Streamlined manual for Verbal A and B (Rough draft). Bensenville, IL: Scholastic Testing Service.

Torrance, E. P., & Mourad, S. (1978). Self-directed learning readiness skills of gifted students and their relationships to thinking creatively about the future. Gifted Child Quarterly, 22, 180-186.

Torrance, E. P., & Safter, H. T. (1986). Are children becoming more creative? Journal of Creative Behavior, 20, 1-13.

Torrance, J. P., Torrance, E. P., & Crabbe, A. B. (1980-1981). Handbook for training Future Problem Solving teams. (rev. ed.). Lincoln: Future Problem Solving.

Treffinger, D. J. (1985). Review of the Torrance Tests of Creative Thinking. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook: Vol. II (pp. 1632-34). Lincoln, NE: University of Nebraska-Lincoln, Buros Institute of Mental Measurements.

Treffinger, D. J., Isakson, S. G., & Firestien, R. L. (1983). Theoretical perspectives on creative learning and its facilitation: An overview. Journal of Creative Behavior, 17, 9-17.

Treffinger, D. J., & Poggio, J. P. (1972). Needed research on the measurement of creativity. Journal of Creative Behavior, 6, 253-267.

Treffinger, D. J. & Ripple, R. E. (1969). Developing creative problem solving abilities and related attitudes through programmed instruction. Journal of Creative Behavior, 3, 105-110.

Vernon, P. E. (1970). Introduction to introspective materials. In P. E. Vernon (Ed.), Creativity: Selected readings (pp. 9-14). Harmondsworth, England.

Wallach, M. A. & Wing, C. W. (1969). The talented student: A validation of the creativity-inteligence distinction. New York: Holt, Rinehart and Winston.

Wallas, G. (1926). The art of thought. London: C. A. Watts.

Webb, J. T., Meckstroth, E. A., & Tolan, S. S. (1982). Guiding the gifted child. Columbus, OH: Psychology.

Wechsler, S. (1981). Identifying creative strengths in the responses to the verbal forms of the Torrance Tests of Creative Thinking. (Doctoral dissertation, University of Georgia, 1981). Dissertation Abstracts International, 1983, 42, 3521A. (University Microfilms Order No. 82-01588)

Welsh, G. S. (1975). Creativity and intelligence: A personality approach. Chapel Hill, NC: University of North Carolina, Institute for Research in Social Science.

Whitmore, J. R. (1980). Giftedness, conflict and underachievement. Boston: Allyn and Bacon.

Williams, F. E. (1976). Rediscovering the fourth-grade slump in a study of children's self-concept. Journal of Creative Behavior, 10, 15-28.

Winne, P. H., Marx, R. W., & Taylor, T. D. (1977). A multitrait-multimethod study of three self-concept inventories. Child Development, 48, 893-901.

Wright, R. J., Fox, M., & Noppe, L. (1975). The interrelationship of creativity and creative self-concept. Psychology, 12(2), 11-15.

Zuckerman, H. (1977). Scientific elite: Nobel laureates in the United States. New York: Free Press.

APPENDIXES

APPENDIX A

DOCUMENTATION OF PERSONAL COMMUNICATIONS

DOCUMENTATION OF

PERSONAL COMMUNICATIONS

Oliver F. Anderhalter, Vice President
Director, Research and Development Division
Scholastic Testing Service, Earth City, Missouri
March 24, 1988; June 9, 1988; June 15, 1988

Orlow E. Ball, Co-author with E. Paul Torrance
Streamlined Scoring Manuals and Workbooks for the Figural (1984) and Verbal (in press) Torrance Tests of Creative Thinking
School Psychologist, Augusta Public Schools, Augusta, Georgia
August 19, 1985; October, 1985; June, 1986; July 16, 1986; January, 1987; February 21-22, 1987; July, 1987; August, 1987; October, 1987; June 28, 1988

Kay Sather Bull, Professor
Applied Behavioral Studies in Education
Oklahoma State University, Stillwater, Oklahoma
June, 1983 - December, 1988
(Committee Chairman)

N. Ruth Christy, President-Elect, OK-OM
Enrichment Teacher
Owasso Public Schools, Owasso, Oklahoma
March 23, 1988; April 16, 1988; July 14, 1988

Anne Crabbe, Director
Future Problem Solving Program
St. Andrew's College, Laurinburg, North Carolina
Chair, Creativity Committee
National Association of Gifted Children
June, 1985; February 18-19, 1987; March 12, 1987; May 30, 1987; November 10-11, 1987; July 7, 1988

Virginia C. Crandall, Fels Professor in Psychiatry
Chief, Section of Developmental Psychology
Wright State University, Dayton, Ohio
July 7, 1985; July 17, 1985; September 5, 1985

Cathy Kass, Associate Professor
Director, Gifted Education
Oklahoma City University, Oklahoma City, Oklahoma
July, 1985

John D. Kauffman, Vice President, Marketing
Scholastic Testing Service, Bensenville, Illinois
February 20, 1987; March 12, 1987; March 16, 1987; May 18, 1987; November 11, 1987; January 8, 1988

C. Samuel Micklus, OM Association, Inc.
Professor, Glassboro State College, Glassboro, New Jersey
Co-founder of Odyssey of the Mind
November, 1983; September 24, 1985; November 10-11, 1987; February 18-19, 1987

Albert I. Oliver, Professor Emeritus
University of Pennsylvania, Philadelphia, Pennsylvania
Member, Board of Directors, OM Association, Inc.
September 17, 1985; November 10-11, 1987

H. Tammy Safter, Assistant Professor
Department of Educational Psychology, Research and Measurement
University of Georgia, Athens, Georgia
October 4, 1985; December 5, 1987

Pauline S. Sears, Professor Emeritus
Stanford University, Menlo Park, California
July 10, 1984; July 10, 1985

Gordon Shewach, Senior Evaluator
Future Problem Solving Program
School of Education
University of Michigan, Ann Arbor, Michigan
March 12, 1987

Edward J. Stehlin, Operations Manager
Research and Development Division
Scholastic Testing Service
Earth City, Missouri
June 18, 1987

E. Paul Torrance
Author of Torrance Tests of Creative Thinking
Founder of the Future Problem Solving Program
Georgia Studies of Creative Behavior, Athens, Georgia
July, 1985; August 15, 1985; November 18, 1987

Joanne R. Whitmore
Assistant Dean for Teacher Education
Kent State University, Kent, Ohio
August 15, 1985; February 26, 1988

APPENDIX B

PROCEDURAL ISSUES: GENERAL CONSIDERATIONS IN THE USE OF MULTIVARIATE ANALYSES

PROCEDURAL ISSUES: GENERAL CONSIDERATIONS IN THE USE OF MULTIVARIATE ANALYSES

Rationale

Utilization of multivariate rather than univariate methods of analysis is preferable in a study of complex related behaviors. Multivariate analysis permits the researcher to measure and simultaneously examine the effect of an independent variable in two or more aspects, dependent variables, of the complex behavior. The use of the single multivariate analysis provides a mechanism to investigate multiple measurements without increasing the possibility that significant differences are due only to chance measurement error with each additional variable (Tabachnick & Fidell, 1983).

The results of any ANOVA are tested against the likelihood that one single obtained effect of that size is significantly different from a chance occurrence. Each additional related measurement taken from the same sample when using the univariate approach adds another hypothesis to be tested and thereby contributes to inflation of the alpha level of the experiment (Larrabee, 1982; McCall & Appelbaum, 1973; Tabachnick & Fidell, 1983).

There are two major alternatives to protect the researcher from inflated possibilities that significant differences are due to chance. The researcher could set the alpha level to test for significance at the appropriately more stringent level, such as .01 or less, or the researcher could use multivariate statistics. The latter is preferred because examination of hypotheses at the stringent level increases the probability of Type I error, missing of important real differences (Larrabee, 1982). Moreover, if the alpha level were to be set at the more stringent level of .01, then the power of the experiment to appropriately locate real differences would be reduced. That is, with an estimated strength of a planned experimental treatment, and known sample size, the researcher can estimate the probability that any statistically significant findings of the study will reflect reality rather than be a chance or error effect. Given an alpha of .05, researchers can increase power by increasing sample size, exerting greater experimental control, or designing the research study so the treatment effect will be clearer (Tabachnick & Fidell, 1983).

Awareness of the problems of inflating the alpha level of a study by use of multiple univariate analyses is relatively recent in the counseling literature (Leary & Altmaier, 1980) and is almost nonexistent in the educational literature. The astute reader in the literature on creativity and on self-concept would, then, note that the studies of creativity typically use a separate univariate analysis of the effects on each dependent variable rather than a single multivariate analysis in a battery (Mansfield et al., 1978). Some studies are suspect to an even greater inflation of the alpha level, by utilizing multiple scores resulting from each separate subtest of the scale as a separate analysis (Casler, 1982/1983), rather than utilizing a total fluency and/or originality score combined from the various subtests as is normally done in studies using the TTCT. Of the studies cited in the present review of the literature which used the TTCT or other instruments which yield multiple scores, only a few of the studies used multivariate analyses of the data (Bolen & Torrance, 1978; Bradley & Gaa, 1977; Kogan & Pankove, 1974; Kolloff & Feldhusen, 1984; Rosenthal et al., 1983; Shaffer-Zamari, 1983; Tetenbaum & Houtz, 1978).

Despite their increased complexity there are several benefits to the use of the multivariate statistical procedures. In multivariate analysis of variance (MANOVA), the primary procedure used in this study, the analyses test the effects

1. Dependence of the variables in the multivariate construct such that the intercorrelation error matrix bears some values over .3 and none greater than .8

2. Logical sense to the construct

3. Homoscedasticity

4. Lack of evidence of multicollinearity or singularity

The major analyses for this study are mixed model MANOVAs, MANOVA procedures in which there are at least two independent variables, one which assesses differences between independent groups of subjects, and another (the repeated measure) which compares differences within the same subjects on different occasions or conditions. In addition to the four basic multivariate assumptions, the use of mixed model MANOVAs requires:

5. Random selection of subjects

6. Normal distribution of the dependent variable

7. Homogeneity of variance and covariance matrix on the dependent variables

8. Independence of subjects. Subjects in one group are unrelated to those in other groups

9. Dependence or relatedness of subjects on the different levels of the repeated factor

All of these assumptions with the exception of number 9 apply to the model for analysis of covariance which also assumes:

10. Linearity

11. Reliability of covariate ($r_{yy} < .8$)

12. Homogeneity of regression

Multivariate analyses of variance examine the effect on a construct rather than on a unidimensional variable. The construct, composed of two or more measures, must show statistical and logical justification that those measures ought to be considered together (Finn & Mattsson, 1978). Statistical justification is provided by examination of the error correlation matrix. One of the dependent variables entered into a MANOVA must bear a correlation with at least one other variable at greater than .3 in order to show statistical confirmation that a multivariate construct has been formed. Any variable with a correlation greater than .8 with any other variable should be eliminated from the construct because it shares too much overlapped variance with the other measure contributing to unnecessary duplication of measurement and, more seriously to the statistical computations, contributing to instability of the error correlation matrix.

The logical evidence for inclusion of a variable into the construct is determined by establishing justification from the literature (Finn & Mattsson, 1978). The prioritization of the dependent variables in the order into the equation is also justified by establishing from prior studies the closeness of the relationship between that dependent variable and the independent variable.

Calculation of Strength of Association: N^2

The appropriate strength of association measure for an ANOVA or MANOVA in which there is unequal _n_ in the cells is eta^2 (N^2). This strength of association measure expresses the amount of relationship, or the proportion of the variance in the dependent variable that is explained by that independent variable (Tabachnick & Fidell, 1983). Two formulas were used in this study to determine N^2. For the multivariate effects, $N^2 = 1 - \Lambda$, where Λ is the Wilks' Lambda statistic. To test the strength of association of univariate effects for the mixed model analyses: $N^2 = \frac{SS_{hypothesis}}{SS_{total}}$, or, for example,

$$N^2_{Similes} = \frac{SS_{OM \times Time}}{SS_{OM} + SS_{error\ between} + SS_{Time} + SS_{error\ within}}.$$

This formula gives the amount of variance in the construct, for example, of creativity which is explained by the dependent variable under consideration, e.g., Similes, when a specific hypothesis is tested, e.g., the OM x Time interaction. It should be noted that in research designs with unequal n in the cells, such as in the present study, it is possible that the N^2 statistics for all of the dependent variables may sum to greater than 1.00 for any given hypothesis.

Use of Dummy Coding of Variables in Regression Analyses

Analysis by multiple regression requires all variables to be continuous or dichotomous. Variables which are nominal, or discrete may be changed into dichotomous data by a procedure known as dummy variable coding (Cohen & Cohen, 1983; Tabachnick & Fidell, 1983). In order to assess the contribution of differences in group membership of OM, grade, and/or teacher in the prediction equations by regression analysis, dummy variables were created for the OM, grade, and the teacher categorical variables. The entrance of the dummy variables was forced into the equation as a set in order to analyze the variance due to the variable, and to permit examination of the effects of the component dichotomous groups (Tabachnick & Fidell, 1983).

Test of the Assumptions for MANOVA and MANCOVA

The reliability coefficients of the dependent variables were discussed in Chapter III and were found to be adequate for a research study. The covariate measures of experience, experience in the enrichment program and prior experience in OM, are measured in fixed categories of time, months or years of previous experience and are thus likely to be sufficiently reliable. Sex and grade level can safely be assumed to be reliably measured and, therefore, also meet the criterion of reliable measurement of the covariates (Tabachnick & Fidell, 1983).

Assumptions of linearity and normality were checked by use of regression analyses using each dependent variable and covariate as an independent variable for a step-wise regression analysis. Regression analyses showed all variables to meet the assumptions of linearity and normality.

The log of the determinant of the within cells correlation matrices was found through SPSS MANOVA for the various analyses to range from: -.17, F(max) criterion = 2.73, df = 3,113 for the repeated MANOVA for OM x Time of three creative dependent variables with scores set at $\leq$ 160; to -6.25, F(Max) = 74.58, df = 10,113 for OM x Time for ten creative creative dependent variables with no transformation of the scores; or to -.82, F(Max) = 199.05, df = 4,108 for the MANOVA of four affective dependent variables; all significant at .000, to -.04, F(Max) = 2.14, df = 3,108, $p > .20$ for the MANCOVA for three creative dependent variables.

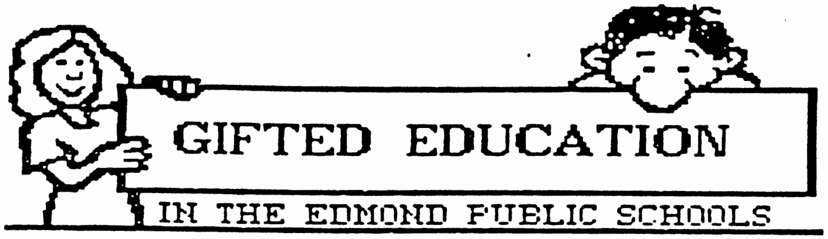
These scores suggest that although each of the individual univariate dependent variables met the assumptions of normality, the tests of sphericity of all of the above analyses, except the creative MANCOVA, suggest departures from multivariate normality. However, "departures from multivariate normality generally have only very slight effects on the Type I error rates . . . As in ANOVA departures from normality may reduce statistical power" (Bray & Maxwell, 1985, p. 33).

When multivariate normality is possibly suspect, the use of transformation of the data is suggested. Two transformations were made of the creativity data, setting the ceiling of standard scores at greater than or equal to three standard deviations above the norm of 100, i.e., scores of greater than or equal to 160, and reducing the ten dependent variables to three by use of the mean verbal and figural Creativity Index measures. These transformations reduced the differences in variability between groups and reduced the number of dependent variables in the study to only a few. Greater confidence can be placed in results of those analyses in which the data was transformed by these two procedures than in the analyses of ten dependent variables. Further reduction of variability, moreover, by means of MANCOVA brought the F(Max) statistic within an acceptable range. Even when there is unequal N, when there are about 20 cases in the smallest group then robustness should be ensured "with a few" dependent variables (Tabachnick & Fidell, 1983, p. 232).

Homogeneity of the variance-covariance matrices was tested by means of Box's M test which is readily available but not very useful because it is known for its extreme sensitivity to departures from normality (Bray & Maxwell, 1985; Tabachnick & Fidell, 1983). Tests of equality of the covariance matrix showed all 3x2 MANOVA designs and MANCOVA designs to meet the assumption. The 3x2x2 designs, however, such as OM x Time x Experience in the enrichment program, with ten creative dependent variables showed Box's M test to indicate failure to meet the assumption, $X^2 = 570.22$, $p < .000$, df=420. When Box's M test leads to rejection at $p < .001$, and sample sizes are unequal, as is the case in the present study, then robustness is not guaranteed (Tabachnick & Fidell, 1983).

If neither condition is met, as in the creativity analyses, and with an increasing number of dependent variables and discrepancy between sample sizes, then a greater distortion of α levels may occur. To evaluate the significance of the global effects, use of Pillais' criterion rather than Wilks' Lambda is recommended to perhaps improve the robustness of the test (Tabachnick & Fidell, 1983). Moreover, if the cells with the smaller samples, rather than the larger cells, are producing the larger variances and covariances, the test is too liberal. Under such conditions, results which support the null hypothesis may be viewed with confidence, but results which indicate significant differences between means may be suspect (Tabachnick & Fidell, 1983). However, when sample sizes are not equal, if Box's M shows $p > .001$, robustness should still be guaranteed.

To check for multicollinearity, regression analyses examined each dependent variable as an independent variable against all other dependent variables, the independent variables of OM, and the possible error covariates: sex, teacher, years of experience in OM and months of experience in the enrichment program. Examination of the output of the 18 multiple regression analyses showed the six analyses of pretest or posttest verbal TTCT scores to yield adjusted R^2 values greater than .80. Although these values are high, they do not approach .99, which therefore indicates that these dependent variables are not likely to be contributing to multicollinearity or singularity (Tabachnick & Fidell, 1983). Moreover, in that none of the multivariate solutions utilizing these variables were unstable in relation to other analyses, these matrices were judged to be meeting the assumptions relating to multicollinearity.



WHO IS GIFTED?

"Gifted" students are defined as those students identified in grades one through twelve as having demonstrated potential abilities of high performance capability and needing differentiated or accelerated education or services.

HOW IS A STUDENT IDENTIFIED FOR GIFTED EDUCATION IN THE EDMOND SCHOOLS?

Students who meet the following criteria will be eligible to participate in the gifted program:

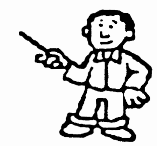
1) Scoring at the 97th percentile or higher on the individually administered Stanford-Binet intelligence test (130 or above).
2) Scoring at the 97th percentile or higher on the individually administered WISC-R intelligence test (129 or above).

The Edmond School provides for testing through the State Department Regional Services. The time for testing once a referral is made is dependent on the schedule of Regional Services. If a parent wishes immediate testing, the Edmond Schools will accept the testing from a private licensed source. The Edmond Schools cannot, however, be responsible for payment for private testing.

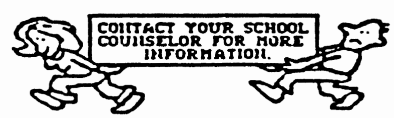
WHO MAY MAKE A REFERRAL FOR THE STUDENT TO BE TESTED FOR THE GIFTED PROGRAM?

Parents, teachers, counselors, administrators, and students themselves (self-referral), may make a referral for testing. Once a referral is made, the school must secure parents' written permission to test. Parents of first grade students who are identified as gifted are encouraged to place children in the regular classroom during the first grade to develop social and emotional skills for positive school adjustment.

WHAT IS THE GIFTED PROGRAM?



The first through fifth grade identified students meet with a designated teacher in a "pull-out program" for enrichment activities during the regular school day. Identified sixth and seventh grade students may elect to enroll in a course for gifted students. Eighth through twelfth grade students may enroll in honors classes. Twelfth grade also offers Advanced Placement classes.



EDMOND PUBLIC SCHOOLS
ENRICHMENT PROGRAM
PHILOSOPHY

The educational philosophy of the Edmond Public Schools is for gifted children the same as it is for every child in the district: to develop his or her abilities to the optimum.

The Edmond Public Schools believes that intellectually gifted children need association with all children to develop a realistic view of themselves and the world in which they live. They need appropriate challenge in their daily school program in the content areas. They also need some opportunity to meet with other highly able children for mental stimulation, motivation, and sharing of interests. As a supplement to their educational program, a one-half day per week "pull out" program is provided identified elementary children. This program presents students with appropriate and differentiated learning experiences; its goal is to help them become better critical and creative thinkers, decision makers, and leaders. The curriculum is based upon the characteristics and needs of most mentally gifted students and is process, not content, oriented.

Edmond Public Schools
Elementary Enrichment Program

STATEMENT OF GOALS AND DESCRIPTION OF PROGRAM

The goals of the Enrichment Program for the mentally gifted are to develop the student's ability to

THINK CRITICALLY	EMPLOY RESEARCH SKILLS
THINK CREATIVELY	EXPLORE, STUDY, AND SHARE AREAS OF INTEREST
FEEL POSITIVE ABOUT ONESELF AND OTHERS	COMMUNICATE EFFECTIVELY

In a broad sense, the goals of the Enrichment Program are to provide differentiated and challenging learning experiences, and the opportunity for stimulating interaction among mental peers, which will enhance the possibility of the gifted to more fully develop their potential and become responsible, caring, and productive citizens and leaders.

STRUCTURE OF THE PROGRAM

As a supplement to their educational program, students in the Enrichment Program meet one-half day per week in a "pull out" program. Each week students are expected to work in learning activities based on the goals of the program. These will be accomplished by the total class, small groups, or on an individual basis. In addition to the work done in class, some out-of-class work will be expected.

It is the philosophy of the Enrichment Program that the gifted student should have an opportunity to make choices and decisions based on self-knowledge of needs and interests. These choices help facilitate individualization and the goal that each child grows toward becoming a self-directed learner. In fostering self-directed learning, students reach toward specific objectives on each task and in an acceptable time frame. They learn to keep records, evaluate their own work, and assume responsibility for their choices and decisions. Students are evaluated each nine weeks by the enrichment teachers on the work and products accomplished in the Enrichment Program.

CONTRACT OPTIONS

A method of choosing a plan for each child's individualized portion of the enrichment work is provided by the contract options. A variety of choices for meeting the contract are introduced in class and individually discussed between the teacher and child. The contract for work requiring out-of-class preparation will then be signed and/or renegotiated with parental suggestions.

The basic minimum contact requires four individually assigned critical thinking workbooks and one Independent Research Study (IRS) as the year's work. Students may choose one of a variety of options to alter the basic plan by substituting a different indepth project for two books of the critical thinking component. Alternative options include projects in areas such as the performing arts, science, math, art, and team creative problem solving (OM).

EXPECTATIONS FOR PLANNING AND PACING OF OUT-OF-CLASS WORK

The IRSs or projects will be planned by the child and teacher with a suggested time frame of the plan. The IRS and project (if one is in the contract) give an opportunity for planning and researching of a topic of special interest to the child. The student is expected to create products to communicate the new learning and to share it with other members of the class. IRSs will be evaluated by the presenting student, the class audience, and by the teacher. The student with the most outstanding IRS for each class is invited to share it at the annual Parent-Child Gathering.

The critical thinking component of the Enrichment Porgram requires each child to complete 2 to 4 individualized workbook packets at a competent level. The number of books varies with the number of other projects the child has contracted to do for the year. It is recommended that the child's outside requirements be divided so that at least one portion of the contract is completed each 9 weeks. A comfortable pace for the child would be completion of at least 1 critical thinking workbook, or a project, or the IRS by the 6th week of each 9 week period.

SAMPLE ACTIVITIES TO ACHIEVE GOALS

CRITICAL THINKING

The four highest levels of thinking are considered to be appropriate for emphasis in programs for the gifted. The majority of the work should evoke the processes of application, analysis, synthesis and/or evaluation* to increase skill in careful, deep and/or abstract thinking.

Conservation tasks
Core subject or unit study (example:
archaelogy, architecture,
futuristics, mythology)
Critical thinking discussions
Critical thinking individualized workbooks
Deductive/inductive reasoning
Evaluating
Figural analogies, classifications
and sequences
Future Problem Solving
Independent Research Study (IRS)
Inferences
Introduction to Logic (series)
Learning centers available as a
required center activity: Analogies,
Circle and Table Logic, Crosswords,
Fill-in, Mazes, Mind-Bender Training,
Syllogisms, Word Search)
Mindbenders
OM (formerly Olympics of the Mind)
Venn Diagrams
Verbal analogies, classifications
and sequences

CREATIVE THINKING

Many activities are offered to increase fluency, flexibility, elaboration, and/or originality* of expression, by means of creative thinking papers, brainstorming, and practice in strategies of creative problem solving.

Art activities
Brainstorming
Creative Problem Solving teamwork
Creating puzzles and other products
Creative writing
Divergent thinking exercises
Droodles
Figural transformations
Future Problem Solving
Hidden figures
Imagin-action
Independent Research Study (IRS)
Junk Sculpture
OM (formerly Olympics of the Mind)
Scamper
Spontaneous problems
Way-out words
Word transformations
Word trees
Wordles

FEELING POSITIVE ABOUT ONESELF AND OTHERS

The affective* development of the child is important to the realization of any person's full potential and is vital in the education of gifted children. In an environment which the children often perceive to be accepting and noncritical, a positive attitude toward learning, a healthy self-concept, and consideration of the rights and needs of others are enhanced. Pride in the gifts and talents of self and others, a sense of responsibility to self and society, and increasing skill and confidence in becoming a self-directed learner, group member, and leader may also be facilitated through the following activities.

*These terms are defined on p. 4.

Book discussions
Boundary breaker and encounter activities
Cooperative teamwork in group Creative Problem Solving
Engaging in work that is motivating
Experiencing success in creating ideas, products, and solutions and in solving complex problems
Listening to others
Opportunities to develop leadership
Problem solving dilemmas
Sharing of work and ideas
Word web
Working together in small groups

RESEARCH

Each student works with the teacher to help plan an Independent Research Study (IRS) which will be presented to the class. Skills in research are introduced.

Bibliographies
Card catalogue
Dictionary skills
Introduction to resources for gathering information such as the almanac, atlas, encyclopedia, interviews, maps, pamphlets, and specialized dictionaries
Note taking
Organizing information
Planning reports
Preparing and presenting an IRS

EXPLORE, STUDY, AND SHARE AREAS OF INTEREST

Opportunities are often available to select materials from a variety of choices for independent or small group work. Work done on independently prepared in-depth studies (IRSs) or on projects will be shared with the class.

Learning centers:

Attributes
Creative writing
Geoboards
Logic Box
Mazes and Illusions
Mini-centers
Origami
Pentominos
Probability
Sprouts
Tangrams
Think Lab
Visual Thinking

Games:

Boggle
Chess
Impuzzables
Mancala
Mastermind
Othello
Pente
Pyramid Puzzle
Tuf
Wordles

EFFECTIVE COMMUNICATION (ORAL AND WRITTEN)

Clarity and organization as well as elaboration and originality contribute to feeling competent in one's ability to communicate well with others. Competency, at grade level, is expected in language skills, spelling, and legibility of handwriting.

Creative Problem Solving:
brainstorming,
offering ideas, listening,
writing ideas, making
decisions, evaluating
Helping to present skits or puppet shows written by other students
Leading a discussion group
Nonverbal communication activities
Presenting oral reports for IRSs or projects
Written work, factual and/or creative

DEFINITIONS

CRITICAL THINKING COMPONENTS OF BLOOM'S TAXONOMY

Application	The ability to use learned material in new and concrete situations. This may also include the use of rules, methods,concepts and principles.
Analysis	The ability to break down material into its component parts so that its organizational structure, or relationship between parts, may be understood. This may also include identification of parts.
Synthesis	The ability to put parts together to form a new whole. Learning outcomes in this area stress creative behaviors, with major emphasis on the formulation of new patterns, structures, plans, communication.
Evaluation	The ability to judge the value of material (statement, novel, poem, research report, art, music) for a given purpose. The judgments are to be based on definite criteria, either internal or external. Learning outcomes in this area are highest in the cognitive hierarchy because they contain elements of all the other categories, plus conscious value judgments based on clearly defined criteria.

CREATIVE THINKING COMPONENTS

Fluency	Generation of a quantity of ideas; flow of thought; number of relevant responses.
Flexibility	Variety of kinds of ideas; ability to shift categories; detours in direction of thought; ability to take different approaches.
Elaboration	Embellish upon an idea; embroider upon a simple idea or response to make it more elegant; stretch or expand upon things or ideas; add details.
Originality	Unusual responses; clever or unique ideas; production away from the obvious; ability to think in novel ways.
AFFECTIVE COMPONENTS	Pertains to the feelings and emotional qualities of a person. Awareness and maturity of one's emotion, feelings, attitudes and values are essential to maximizing the learning process and to full functioning as a person.

The following components of creative thinking are considered to enhance the affective qualities of creative behavior:

Complexity	Imagination
Curiosity	Risk Taking

RESPONSIBILITIES OF ENRICHMENT PROGRAM STUDENTS

Students are expected to:

1. Check with their teachers each week to see what work they need to make up.

2. Demonstrate competency in their basic grade level academic and developmental tasks.

3. Bring a folder with a loose-leaf paper and a pencil to class. Such a folder will help students to carry notes, a current thinking workbook, and other assignments.

4. Help their parents keep informed by delivering communications.

5. Record and evaluate work in class each week.

6. Learn to think for themselves and do their own work. Encouragement, support, and thoughtful questioning from parents is helpful.

7. Demonstrate appropriate classroom self-discipline.

8. Develop skills in independent and/or small group work to become a self-directed learner.

9. Take care of class materials and assist in their storage.

10. Attend class for the designated time, unless there are special circumstances which have been confirmed by parents or school professionals.

11. Be productive in Enrichment and complete required class assignments. Select and complete at least four required centers and several optional centers.

12. Complete an individualized contract for out-of-class work. The contract will include an independent research study, some individualized critical thinking workbooks and possibly an additional project. Frequently students will be asked to rethink problems in order to earn a final grade of 82% or higher for credit on their workbooks.

13. Demonstrate an interest in and/or a commitment to the program.

APPENDIX D

INFORMATION ABOUT ODYSSEY OF THE MIND

Edmond Public Schools
Enrichment Program
October 21, 1985

Dear Students and Parents,

Last year 130 students, 1st through 10th grade, participated in the Edmond OM Program. The Seventh Annual Edmond OM competition will be held on Saturday, March 1, 1986, at the Special Services Center. The Spirit of the Problem for each of the 1986 OM competition problems and the non-competitive primary problem are being distributed at the present time.

The full problems will be read and discussed in Enrichment classes as soon as possible. Students will brainstorm the problems and see slide shows of last year's competitions. OM (formerly Olympics of the Mind) is a voluntary activity in which Enrichment rogram students are eligible to participate.

A note about outside assistance. The National and Oklahoma OM Associations, the Edmond Public Schools, the Edmond OM Parent Advisory Committee, and your Enrichment Program teachers sponsor the OM program for you. Our goals are to have students learn (a) techniques for solving problems, (b) how to follow through with ideas and bring them to a completed solution, (c) how to work with others in a productive manner, and (d) to appreciate sportsmanship. Non-team members, including parents and coaches, are <u>not</u> allowed to assist the team in solving the problem.

Each team must have a coach who attends a two-hour coaches' workshop and is responsible for training the team under the general guidance of the Enrichment teacher. The coach is usually a parent, but other interested trained volunteers may serve. A conservative estimate of the time involvement for the parent coach is 30 hours prior to the local competition. Sharing of coaching duties as co-coaches can often work well. At least one coach per team, even if previously trained, must also attend one of the two-hour coaches' workshops.

"The coach is not to work on the solution to any problem." The coach may tell the team that he/she likes or dislikes the solution and may encourage more brainstorming for solutions or require further refinement before accepting the solution for a competition. She should also encourage the inclusion of style and assist in obtaining materials or facilities to work on the problems. A coach may arrange for guest speakers, field trips, etc. To provide the team with a working knowledge of problem related materials, the resource people should not be familiar with the specifications of the problem. "Teams must design and produce their own problem solutions (including costumes)" OM Coaches' Manual, 1984.

When teams are formed the adult coach and all team members must certify that they understand the rules about outside assistance and that only the team members must solve the problems. In order that students have sufficient time for their work, teams may begin forming now; however, the final composition of a team is the responsibility of the teacher contact coach in cooperation with the team's coach. A team may start meeting <u>after</u> a coach or co-coach has attended the coaches' training.

When your child brings home the full problems which describe the limitations, specifications, rules, of play, and scoring points, please follow these guidelines: Please <u>do not</u> make any suggestions as to how the problem may be solved, or the kind of tools or materials which could be used in construction to solve any aspect of the problem. Such suggestions would be classified as illegal outside assistance and could lead to a 50-point penalty against the team's score. Please <u>do</u> read and discuss the problems with your child. As you do so, please feel free to help your child to think of her own ideas, define terms with which she is unfamiliar, and encourage him to expand his own thinking through imagination.

We do not require participation on an OM team but do strongly encourage it. Whether the child's team wins or not, we find that many of the students enjoy the work, feel very productive, and gain from the teamwork and creativity experiences. The work is done in teams of 3 to 5 members at home under the guidance of a trained volunteer coach. According to new national rules, it is strongly recommended that teams which participate at the state or world's finals level have 5 members with 2 additional alternates. If your child is considering participating in OM this year, please consider the possibility of assisting the team as a parent coach or assistant coach. We also welcome new members on the Parent Advisory Committee to share representation from each school. The committee meets once a month on Friday afternoons. The next meeting will be Friday, November 8, at 1:30 p.m., Special Services.

<u>Dates</u>: Parent Coaches' Training--Room 108, Special Services, 215 N. Blvd.; Tuesday, November 19, 6:30-8:30; Thursday, December 5, 6:30-8:30; Monday, January 13, 6:30-8:30.

Deadlines for Team Registration and payment of $5.00 fee:	January 24, 1986
Annual OM Orientation	October 29, 7:00 p.m.
Edmond OM Competition	March 1, 1986 (Saturday)
Regional Competition	March 15, 1986 (tentative)
State Competition	April 12, 1986 (Ada)
World Finals	June (Arizona)

Sincerely yours,
Anne Fishkin
Kathleen Kaufman
Barbara Sharp

Anne Fishkin, Kathleen Kaufman, Barbara Sharp and Janice Dillon, OM Contact Coaches and OM Parent Advisory Committee

OK-OM

(Formerly Oklahoma Olympics of the Mind)

OM Creative Competitions is presented under the auspices of the national organization: OM Association, Inc.

CREATIVE COMPETITIONS STIMULATE AND TRAIN YOUNG THINKERS

The state creative competition for the OK-OM brings together some of Oklahoma's finest young creative thinkers. Through involvement in OM, students receive hands-on training in creative problem solving which will assist them in later activities in their communities, nation and world.

HISTORY OF CREATIVE COMPETITIONS

"Creative Competitions: An Olympics of the Mind" was initiated in New Jersey in 1978, sponsored jointly by the New Jersey Department of Education and Glassboro State College. Since that time, student and school district participation has grown within the state and across the nation. Each year a national meet is held and schools from across the country send their state winners to compete. Oklahoma was one of the first states west of the Mississippi to be represented in this national competition. Creative Competitions, a private, not-for profit corporation has been recognized by the national Department of Education as fostering and developing creativity. Oklahoma OM is approved by the Oklahoma Secondary Activities Association.

PURPOSE OF OK-OM CREATIVE COMPETITIONS

OM's Creative Competition encourages creative thinking in youth by providing opportunities for students, either individually or together in teams, to solve problems in imaginative, creative ways, with coaching provided by trained adults. The problems require that students create an actual product or solution which they present as their entry. As persons who have attended previous state competitions can testify, OM is one of the most stimulating activities young people can experience. The students become totally absorbed in their project, and best of all the participants know that their solution came from within their own creativity, individually and as a cohesive group.

THE PROBLEMS

Long-term problems, which change every year, are sent to each member upon joining the association. Examples of past problems range from making a vehicle that will run on the energy of five pounds of sand dropping two feet to rewriting a chapter of the Odyssey and performing the rewritten chapter. The long-term problems fit a wide range of interests and subject areas.

Long-term problems are given to members (who form teams) in advance of local, district, state or world competition. This gives the teams time to prepare their solution to the problems. There are specific design specifications and monetary limitations in solving long-term problems. (Maximum score = 200 pts.)

Spontaneous problems are given to the teams on the oay of the competitions to challenge their ability to "think on their feet." The problems require teams to give creative responses to a verbal or hands-on problem. Spontaneous problems are different at each competition. (Maximum score = 100 pts.)

Judges also give consideration to the "style" of the presentation that elaborates their long-term problem solving (i.e. costumes, music, art, etc.). (Maximum score = 50 pts.)

The combination of long-term, spontaneous, and style scores determines the team scores in competitions. (Maximum total score = 350 pts.)

COMPETITION LEVELS

There are three divisions in OM competitions: Division I, Grades K-5; Division II, Grades 6-8; and, Division III, Grades 9-12. There is also a non-competitive problem for grades K-2.

START YOUR OWN OM PROGRAM

Those interested in coaching an OM team may do so by joining the state and national groups and attending a one-day coaches' training workshop, thereby receiving their certification.

See attached application for membership dues.

MEMBERSHIP PRIVILEGES:

National
- A Coaches' Handbook (national edition)
- Six Long-Term Problems
- Membership card
- Subscription to OM Quarterly Newsletter
- Membership also supports the chartered association competitions and the Worlds Finals Competition

State (individual and/or coach)
- Subscription to OK-OM Newsletter
- Voting privileges in state association
- Membership supports the OK-OM charter association and regional and state competitions

Registration for Competition
To register teams for any state-sponsored competition it is necessary to have:
- A trained contact peson who has attended the 5-hour state coaches' training session within the past two years.
- A national membership which covers the team and possible other teams
- An individual state membership for each coach
- Student dues paid by all team members

Competition at the national finals is limited to winning teams from each division at the Oklahoma State competition.

OM BOOKS

Looking for practice problems for your team? Want some diversified problems for your classes? Here they are! **Odyssey of the Mind-Problems to Help Develop Creativity** by C. Samuel Micklus, Ed.D. is available. It includes both long term and spontaneous problems and has a *special section on coaching tips and suggestions*. It was edited by Dr. Al Oliver, Professor Emeritus, University of Pennsylvania and current member of the OM Board of Directtors. If you wish to order, send a check or purchase order for $14.25 (includes postage and handling) to OM Association, Inc., P.O. Box 27, Glassboro, NJ 08028.

Still available is **Problems, Problems, Problems** by Drs. Gourley and Micklus. To order send a check or purchase order for $12.70 to OM Association, Inc., P.O. Box 27, Glassboro, NJ 08028.

A videotape and/or slide show may be rented from OK-OM. For rental fee and schedule please contact the OK-OM Secretary.

Coaches' Training Sessions and other Staff Development Workshops can be arranged. For further information see the attached Staff Development leaflet or contact the OK-OM Secretary.

OK-OM

"Imagination is more important than knowledge for knowledge is limited, whereas imagination embraces the entire world - stimulating progress, giving birth to evolution." —Albert Einstein

Sample Problems from the 1985-1986 OM Competition

Problem 2. Technocrats
The problem is to design, to develop and to mass produce a product. The design and production planning stages will be completed in advance. During the competition the team will produce ten reasonably identical items. Part of the team's score will be derived from the product's design, complexity of the product, and creativity of the production line. The Spirit of the Problem is to develop a product, to develop a system to produce the product, to package it and to move it to the shipping dock.

Problem 3. Great Art Lives
The problem is to select two works from one of the art masters listed and to paint and/or sculpt replicas of these two works. One work will be used as a prop or in the set as a decoration. The second work will be the subject of the team's performance. The Spirit of the Problem is to replicate accurately two of the artist's works, to produce an original work, to represent the artist, and to make a work "come alive." In addition, the team will also design and make a playbill.

Problem 4. Bridging the Gap
The problem is to design two structures which are made of 1/8" x 1/8" strips of balsa wood and glue. A beam structure will be placed onto two 2" x 4" blocks and a load-bearing structure will then be placed onto the beam structure. Weights will then be placed onto the load-bearing structure, one weight at a time, until the structure breaks or the eight minute time limit expires. The Spirit of the Problem is for the team to design and to construct a structural unit which will hold weights.

Problem 5. History . . . The Way It Was
The team is to select one historical event from a given list. A "Master of Ceremonies" (MC) will describe the event as history recorded it, or "the impact that event had on humankind." The MC will then inform the audience of "how it really happened." This portion of the problem will then become the team's humorous interpretation of that same event. The Spirit of the Problem is to select an event, to create a humorous interpretation, to make necessary props and to perform the solution.

APPENDIX E

ISSUES IN MEASUREMENT OF THE TORRANCE TESTS OF CREATIVE THINKING

ISSUES IN MEASUREMENT OF THE TORRANCE TESTS OF CREATIVE THINKING

Divergent Thinking Tests and a Definition of Creativity

As discussed in Chapters I and II (pp. 13-15; 25-31), creativity can be defined from a personality, a process, a product, or an environmental perspective (Arieti, 1976; MacKinnon, 1975). Divergent thinking tests have been considered to be measures of the creative process (Khatena, 1982) or to be measures of the individual's potential ability to behave creatively (Torrance, 1974b). If the intended use of the measure of divergent thinking is for research, rather than for identification to discern among various personalities or to determine who tends to behave creatively, however, we are then looking at a product interpretation of creativity. When looking at behaviors to study the effects of a research treatment condition, responses to divergent thinking tasks may, therefore, appropriately be considered to be products or outcomes of a person's tendency to behave creatively (see definition of creativity, pp. 13-15).

Issues of Validity

Content and Construct Validity

Creative behavior may occur when three necessary components are present: the person has the ability to perform creatively, has the skills to do so, and is motivated to perform creatively (Torrance, 1979). The Torrance Tests of Creative Thinking (TTCT) batteries are primarily conceived to be addressing the first of these components, to be a measure of creative thinking abilities (Torrance, 1974a, 1974b). The TTCT offer a variety of tasks to which the test taker responds in either a verbal or a figural modality. Responses to the activities in both batteries display a varied range of creative behaviors. The resultant scores reflect creative thinking ability without necessarily reflecting artistic skill. This author's experience with these tests show that the scoring procedures produce a score of a student's figural creative thinking abilities that is remarkably uncontaminated by artistic skill. The broad sampling of behaviors elicited by tasks free of technical or subject matter restraints contribute to Torrance's (1974b) view of evidence that the TTCT has appropriate content validity.

Do the standard scoring categories--for the verbal tasks, fluency, flexibility, and originality, and, for the figural tasks, fluency, originality, and elaboration--reflect the broad range of creative behaviors that are exhibited? Torrance presented in the Norms-Technical Manual (1974b) evidence of validity of the standard scoring categories with a variety of constructs of creativity and predictions of a variety of future creative accomplishments (see Chapter II, pp. 35-37). However, although the tests "have proved useful in educational practice, many users have made two major criticisms: the scoring is too time-consuming and the tests assess only the divergent production abilities and do not tap the essence of creativity" (Torrance & Ball, 1984, p. 5).

The development of the streamlined scoring for the figural battery reflects extensive experimentation with a variety of norm-referenced and criterion-referenced indicators. The creative strengths indicators, which were added to the scoring, adequately met the following criteria: they showed up clearly in test behaviors, occurred with adequate frequency, and appeared to be sensitive to effects of training and practice (Torrance & Ball, 1984). The addition of the newer scoring categories of resistance to premature closure and of abstractness of titles, as well as the indicators of creative strengths, probably improve characteristics of the TTCT results as a measure of creative thinking that is relatively unrelated to artistic and/or technical skill. The scoring category of abstractness of titles adds a dimension of creativity to the resultant score that is different from originality, as defined as a statistically infrequent idea. Torrance (1979) considered a high ability in abstractness of titles to indicate strength in the important abstract, synthesizing qualities of creative behavior. This ability taps into the "creativity that comes from the collision of opposites and their integration" (p. 55) which thus reflects the creative ability to "highlight the essence" (p. 55). He considered resistance to premature closure as a primary measure of qualities of psychological openness. Openness is considered to be necessary for adequate incubation of ideas and therefore should be related to skill in deferring judgment when appropriate.

Criterion Validity

Torrance and Ball (1984) cited the validity studies that were presently available for the streamlined measures. The original protocols from a 22-year predictive validity study were rescored and compared with two measures: the ratings of quality of adult achievements and the recognition of the number in the sample who were creative as adults. The total indicators of creative strengths and the Creativity Index of elementary school TTCT performance bore significant correlations with the two indicators of adult creative achievement for all groups. The norm-referenced scores for abstractness of titles and resistance to premature closure correlated at least as well as elaboration, and with greater validity than the originality or fluency scores with the two criteria of adult creativity.

A recent study by Gray (1986) found significant differences in the streamlined TTCT scores to be associated with guided imagery work offered to half of the fourth grade children in a series of creative drama classes. She found that scores for the children who had the guided imagery as well as the creative drama were significantly higher in fluency, originality, and the Creativity Index, but were significantly lower in elaboration, titles, and closure. These findings suggest that the new scoring categories of the streamlined scoring system answer the problem of assessing the essential qualities of creativity (Torrance, 1979) to a far greater extent than use of divergent production measures alone, that is, fluency, flexibility, elaboration, and originality.

Reliability

Divergent thinking instruments designed to assess creativity are open-ended in the range and quality of possible responses; therefore, internal consistency methods of assessing reliability of instruments which yield forced choice, multiple choice, or other forms of fixed-range responses are obviously not appropriate to determine whether the creativity measures are consistent (Treffinger & Poggio, 1972). Three methods of determining reliability are appropriate: 1) alternate forms, 2) test-retest, or stability (Nunnally, 1970), and also 3) inter-rater reliability. The TTCT, like other tests of divergent production, is subject to several sources of unreliability: 1) errors in scoring the tests, 2) errors due to subjectivity of measurement from the rater's standpoint, 3) errors due to the testing environment, and 4) errors due to fluctuations in the individual (Nunnally, 1970). Evidence for establishing reliability from the above methods and some of the related sources of unreliability are discussed below.

Alternate Forms/Test-Retest Reliability

Two studies of counterbalanced design to assess alternate forms reliability were cited for the original scoring (Torrance, 1974b). One study, which had only two weeks between testing sessions, obtained alternate forms coefficients of .84 to .93 for the three verbal scores, and .71 to .85 for the figural scores. The other study apparently had an eight-month time interval between testing sessions and an experimental creative writing treatment for half of the subjects. Alternate forms reliability in the counterbalanced design with a longer time interval between testing sessions yielded equivalency of forms reliability of .61 to .87 for the verbal tests and of .60 to .80 for the figural tests, with one of the eight coefficients, figural fluency at r=.50; not an adequate reliability level for research purposes.

Several test-retest reliability values were cited by Torrance (1974b) from a variety of studies. Some of these used abbreviated test batteries consisting of as few as four activities, two verbal and two figural. Most of these studies administered the alternate forms (A and B) over varied time intervals, ranging from one week to three years. The obtained reliability coefficients ranged from .35 to .89. Many of the reliability coefficients were greater than .70. Evidence of alternate forms or test-retest reliability was not presented in the manual for the streamlined version (Torrance & Ball, 1984).

The level of reliability coefficients would suggest that the TTCT is adequate for research use but, perhaps, at questionable levels of reliability for purposes of identification. When reliability is estimated by the test-retest method which does not use alternate forms, the reliability coefficient is usually an overestimate (Nunnally, 1970). The levels of test-retest reliability are frequently below .80, and TTCT performance is susceptible to testing variables that affect motivation as sources that contribute to test unreliability (see below for example of the context effect). This leads one to question whether the TTCT batteries as a measure of an individual's abilities in creative thinking are a sufficiently dependable determinant for purposes of individual prediction. Further assessment of reliability of the TTCT in its use for identification is found in the discussion below of identification of the creatively talented.

Inter-Rater Reliability

Adequate evidence that high inter-rater reliability can be obtained for the TTCT is presented for both the standard scoring procedures and the streamlined procedures (Torrance, 1974b; Torrance & Ball, 1984). By careful study of the manual, inexperienced scorers can obtain inter-rater reliability with trained scorers at values greater than or equal to .66, with most scores correlating greater than or equal to .88 (Torrance, 1974b).

Confirmation that self-trained scorers can score the TTCT with adequate reliability was found by Halpin and Halpin (1974) and by Rosenthal, DeMers, Stillwell, Graybeal and Zins (1983). Halpin and Halpin found inter-rater reliability coefficients at .92 or greater for all TTCT measures. The scores from two self-trained graduate student raters were highly similar in rank to the scores assigned by a professionally trained scorer. There were, however, significant differences in the values assigned to four of the seven measures. The scores for figural and verbal fluency and for verbal originality received mean values from at least one of the self-trained raters that were significantly different from the values assigned by the professionally trained rater. Of the seven measures, the mean values assigned to figural elaboration were most discrepant; mean values for the 15 test booklets were 94.40 to 96.07 from the self-trained scorers but 122.00 from the professionally trained scorer. Halpin and Halpin recommended that a single scorer should be used for any study, and that local norms based on that scorer's work should be developed in preference to use of national norms for self-trained scorers.

Rosenthal et al. (1983) compared inter-rater reliability for fluency and originality for abbreviated batteries of the TTCT. They administered four subtests--from Verbal A, Unusual Uses and Just Suppose, and from Figural B, Picture Completion and Circles--to two groups of elementary aged children, nongifted and gifted. The gifted children participated in the district's gifted program and were selected by demonstrating group IQ scores over 120 and reading and math achievement over the 77th percentile. The average IQ for the random sample of 25 gifted children was 131, and the average IQ of the 25 children who were not in the gifted program was 114.

Because gifted students were excluded from the standardization sample of the TTCT and only the mid-range representation of most school populations was used as the norms comparison group for the TTCT (Rosenthal et al., 1983; Torrance, 1974b), Rosenthal et al. hypothesized:

> . . . that the inter-rater reliability coefficients would be lower for the gifted students, since they would produce more divergent or unusual responses, such that the scoring manual would not be a sufficient guide for this population. Also it was hypothesized that mean differences in the creativity variables generated by the TTCT would appear across groups; across raters, as in the Halpin and Halpin (1974) study; and that these differences would be most pronounced in the gifted sample. These results, if found, would further support the need for specific scoring procedures for the gifted population (p. 36).

Three self-trained scorers independently scored the four TTCT subtests for fluency and originality. Inter-rater reliability coefficients for verbal fluency were $r \geq .85$, figural fluency $r \geq .93$, verbal originality $r \geq .65$, and figural originality $r \geq .77$. The inter-rater reliability judgments, however, unlike the hypothesized relationship, were more consistent for the gifted group than the nongifted group: all $r \geq .85$ in the gifted group. Comparison of the effects of the two different independent variables, group and rater, was appropriately tested by multivariate analysis of variance (MANOVA) (see definition Chapter I, and Appendix B). The MANOVA yielded a statistically significant difference between raters but no significant differences were found for effects involving group as gifted or nongifted. Although the inter-rater reliability coefficients were high, there were significant differences between raters in the level of their scoring for all four variables. That is, there was consistency between raters in their respective ordering of the subjects within the sample, but there were differences between raters in the values they assigned to the scores. Rosenthal et al. (1983) concluded, based on their findings and the similar results by Halpin and Halpin (1974), that it is important to use a single rater in "scoring for a particular population of students, particularly where cutoff scores are used" as a basis for selection for a gifted program (Rosenthal, et al., 1983, p. 40). Although not mentioned as a recommendation by Torrance in his Guidelines for Administration and Scoring Comments on Using the Torrance Tests of Creative Thinking (1987), personal communication from J. D. Kauffman (November 11, 1987, see Appendix A) corroborated the importance of using a single scorer in a given research study or identification search.

Sensitivity of the Tests to the Testing Environment: A Source of Unreliability or an Issue in Validity?

It is a myth to consider that measurement of a person's creative thinking abilities is some absolute measure of that person's native creativity (Khatena, 1982). This author agrees and considers that the sensitivity of creativity tests to the testing environment (Elkind, Deblinger & Adler, 1970; Torrance, 1987) and to motivational variations in the administration of the battery (Torrance, 1972a, 1987) contributes to the validity of tests of divergent thinking as a measure of an individual's potential creative thinking ability.

Effects of Warm-Up

As stated above, to observe creative behavior, it is essential that motivation to produce creative energy should be aroused (Torrance, 1979). A survey of studies which varied warm-up activities, for example, showed that, in general, brief "psychological warm-up given before testing results in small but consistent and statistically significant gains over standard conditions" (Torrance, 1987, p. 12). Moreover, the type of warm-up chosen should be designed to help activate the type of thinking to be tested; that is, some kind of figural activity would be more appropriate as a warm-up to a figural test than to a verbal test (personal communication, C. Kass, 1985, see Appendix A).

Effects of the Environment: The Context Effect

Sensitivity of the tests to the testing environment is known as the context effect. From several studies surveyed by Torrance (1972a, 1987) which varied the testing environment, it was clear that results on the TTCT and other measures of divergent thinking can often reflect differences in testing conditions, such as variation in encouragement, reward, cue-rich environment, or type of activity children were engaged in prior to testing. The findings of two of these studies are very pertinent to the present study. In a study by Mohan (1970, cited in Torrance, 1972a) fourth grade students were tested in a cue-rich or a cue-poor testing room. High creative children showed greater responsiveness on their TTCT scores when tested under cue-rich conditions than did low creative children.

Elkind et al. (1970) studied creativity in children attending a school which offered "a non-graded flexible curriculum . . . in a highly engaging setting" (Elkind et al., p. 351). A prior evaluation study of the school showed a surprising finding: children attending the innovative school did less well on the creativity tests than a control, matched group who attended public schools in the same city. The authors suspected that the children enjoyed their activities at the innovative school, in particular their elective interest areas, so much that they were not motivated to perform at their best on the creativity battery which was experienced as an interruption of tasks they enjoyed.

> Many of the public school children appeared reluctant to return to their classroom and seemed to enjoy the testing as a novel and interesting experience. Just the reverse appeared to be true for {the children from the innovative school} who had seemed engrossed in what they were doing and hence reluctant to leave for the testing and eager to return to the classroom (Elkind et al., 1970, p. 352).

The context effect experiment, then, used children at the innovative school as their own controls. They were each administered equivalent forms of three subtests of the Wallach-Kogan battery with the testing session interrupting two kinds of ongoing activity. The testing session for the "interesting" condition interrupted an activity that was noted by the child's teacher to be a favored activity. The "uninteresting" condition was a repetitive test of study skills that was contrived by the experimenters. The tests were administered in a counterbalanced design for order of equivalent form and for order of type of interrupted activity. The results showed very marked differences in tested creativity under the two context conditions; the mean number of responses for children tested when their ongoing activity was "uninteresting" was 57.09, but the mean number of responses given by these same children when they were tested under reluctance to leave a favored activity was only 32.09. Elkind et al. (1970) concluded that scores on some measures that are considered to be testing creativity are considerably affected by the kind of ongoing activity that is interrupted by the testing session.

On the basis of these two studies by Mohan (1970, cited in Torrance, 1972a) and Elkind et al. (1970), one might surmise that children who are highly creative might not do as well on creativity tests when the testing session interrupts favored activities as when the testing session interrupts an ongoing activity that is not intrinsically interesting to them. The failure of users of creativity tests to be sufficiently cognizant of the sensitivity of the instrument to motivational effects such as the context effect can contribute to a loss of reliability, that is, to lessened consistency of measurement for a given individual.

Motivational Effects in this Study

In the present study, the sample consisted of elementary school gifted children who, based on pretest scores on the TTCT, were highly creative. Pretest mean verbal scores were 132.41 for the sample of 116, greater than one and one-half standard deviations above the norm of 100, with 17 children obtaining mean verbal scores ranging from 160 to 229, from three to six standard deviations above the mean. All group administered tests and most of the individually administered verbal tests for the younger children were given during the children's enrichment class, which meets for only one-half day per week. Some of the second and third grade children were tested for the verbal battery in after-school appointments rather than during their enrichment class times. Standardized tests are not administered during enrichment classes with the exception of the year of the study.

Most of the children in the enrichment program are usually highly engaged, often in a self-directed manner, in their activities in the program. Most seemed quite willing to engage in the pretest activities. Although many of the tasks in the TTCT battery are similar to some creativity training activities used in the curriculum, the pretesting appeared to be perceived by the children as a novel experience. However, in the posttest, many of the children were much less interested, and a few were even uncooperative in their attitude. The verbal TTCT posttests were, even more than the figural TTCT or other measures in the study, perceived as uninteresting tasks compared to the ongoing activities of the enrichment class. Moreover, the total amount of time involved in testing for the study was about 7½ hours for all pre- and posttests, representing more than 1/12 of total time in enrichment classes for many of the children. The three teachers (and three examiners for the individual testing) administered warm-up activities, utilized standard administration procedures, and were careful to call the TTCT tasks activities rather than tests. However, from the remarks of several of the children, it appeared that the motivational level was clearly lower for the verbal posttests than for the pretests.

The findings of Mohan (1970, cited in Torrance, 1972a) and Elkind et al., (1970) then and the lower motivation observed by the teachers during administration of the posttest help to explain the surprising finding of significantly lower posttest scores obtained in the verbal battery for this study (see Chapters IV and V, results and discussion). In hindsight then, this investigater, although cognizant of the context effect research while planning this study, would have been wise to consider the possible motivational influences of overtesting more seriously. In retrospect, then, if abbreviated batteries of the TTCT had been used instead of the full batteries, this study might, like many other studies of creative problem solving (CPS) (Rose & Lin, 1984), have reflected positive results associated with the creativity training of OM and of the enrichment program.

Contributions of this Study to Reliability Information

Alternate Forms Reliability

As noted in Chapter III, test-retest reliability coefficients were computed for all instruments used in this study. Test-retest reliability using the same forms under controlled (no treatment conditions with a short time interval) will usually indicate the upper bound of the reliability because memory of the test items will inflate the correlation between the two times of testing. Alternate forms reliability is therefore a preferable method of estimating reliability of a test (Nunnally, 1970). The reliability coefficients produced in this study, however, are likely to have resulted in underestimation of the test-retest reliability. In this study, posttests were administered five months after the pretests. There was an intervening treatment condition for all the children including those in the non-OM group, that is, all the children experienced a broad variety of challenging activities including Future Problem Solving during their participation in the enrichment program (see Procedures, Chapter III). In addition, two of the groups participated in OM and exerted high or low effort on their teams. Because these treatments produced significant differential effects on some of the 14 dependent variables, the reliability coefficients are to some unknown extent underestimated.

The TTCT reliability coefficients were obtained by correlating the standard scores from the two forms, A as pretests, B as posttests. The raw score values were used for the total of creative strengths. All figural booklets were scored by one scorer, the investigater of this study, within a two-month time period. The verbal booklets were scored by the Scholastic Testing Service (STS) professional scorers. It is possible, but not known, that the verbal A and verbal B booklets may have been scored by different experienced scorers at STS. Although STS conducts stringent and regular inter-rater reliability studies on their scorers to maintain high levels of consistency of scoring (O. F. Anderhalter, personal communication, June 9, 1988), if the verbal A and verbal B were scored by different raters, some additional error variance would be contributed by differences in raters, thus possibly further lowering the alternate forms reliability coefficient for the verbal booklets.

The retest reliability coefficients for the creativity, probably for the creative self-concept, and possibly for the locus of control, measures in this study are likely to be conservative also because of the restricted range of abilities in this sample. All children in the sample are mentally gifted, at individually tested IQ greater or equal to 130. A low positive correlation exists between intelligence and creativity when IQ range is below IQ of 120 (Davis & Rimm, 1985; Taylor, 1975). Therefore, this sample may be considered to be restricted to children with creativity scores in the above average ranges. The reported means (in Table F-9, Appendix F) confirm that the sample is restricted to children with significantly above average creativity; the average posttest scores for the verbal TTCT were greater than or equal to 128.73, for the Creativity Index greater than or equal to 123.21, and for Similes at 43.43 (see also Chapter IV, p. 141-145). Correlations performed on a restricted range of scores are an underestimate of the relationship between the measures, that is, reliability coefficients are larger when based upon more diverse groups (Nunnally, 1970).

These reliability coefficients may also be underestimated because they were calculated on the standard scores rather than the raw scores (see also discussion below on conversion to standard scores and in Chapter V, pp. 209-211). Conversion of the raw scores to the standard scores adds a source of error of measurement by changing the spread of the scores by differing amounts for the Form A and Form B versions of the test. The preferred method to calculate reliability for the TTCT is on the raw scores rather than standard scores (O. F. Anderhalter, personal communication, June 9, 1988).

The means, standard deviations, and Pearson correlation coefficients are reported in Table F-9, Appendix F. The TTCT verbal correlations range from .45 to .51 with the mean verbal, average of the three verbal scores, at $r = .53$. The figural TTCT reliability coefficients, with the exception of resistance to premature closure, range from .32 to .45, with a reliability coefficient for the Creativity Index at $r = .45$. All of these correlations were significant at $p < .01$. The reliability coefficient for closure, however, was only $r = .18$.

The alternate forms reliability coefficients for verbal TTCT scored by standard procedures and for figural scored by streamlined procedures shown in Table F-9 report values that are considerably lower than those cited by Torrance (1974b) for verbal and figural TTCT scored by standard procedures. As noted above, there are several reasons why the reliability coefficients for the verbal TTCT battery and for figural fluency, elaboration, and originality are lower than those coefficients cited by Torrance. That is, in this study: 1) the sample was restricted to children with IQ greater than 130, and therefore could have been restricted to eliminate subjects at the lower ranges of creativity, or there could be no relationship between intelligence and creativity; 2) the study produced a significant intervening treatment effect; 3) the verbal booklets were perhaps scored by different raters; 4) the coefficients are calculated on converted standard scores rather than raw scores, and 5) there appears to have been a significant effect of regression toward the mean from pre- to posttesting. There are no prior alternate forms reliability coefficients given by Torrance and Ball (1984) for the streamlined scoring so there are no comparable values for abstractness of titles, closure, or total of creative strengths.

Given the above reasons for lower obtained reliability coefficients and a consensus of prior reported values at greater than .60 (Torrance, 1974b), it is reasonable to conclude that the reliability coefficients in this study might be underestimated by as much as .20. There are two affective measures in this study for which a comparable test-retest comparison is available, the Sears and the IAR. The obtained test-retest reliabilities in this study are below those reported by Crandall (1978) by approximately .22 for retest agreement after a two-week interval.

However, the Sears retest reliability in this study is higher by .19 than the stability coefficient reported by Sears (1972) for intact, non-experimental classes tested after a four-month interval. There is evidence to show relation between IQ and creativity (Taylor, 1975) and IQ and internalization of locus of control (McKenna, 1981). Therefore, the lower estimates of reliability in this study for the TTCT and IAR may in good part be explained by the restricted ranges of those measures exhibited in this sample of gifted children.

The low reliability from alternate forms on resistance to closure, however, is of concern. The coefficient of .18 indicates inadequate alternate form reliability thus indicating for the closure measure that the forms are not equivalent. The resistance to closure measure, unlike the other five measures for the figural streamlined scoring, is based on responses to only a single subtest, Incomplete Figures. Differences between forms A and B might be more pronounced when only a single subtest is considered without the contribution of scores from other subtests.

Inter-Rater Reliability

An inter-rater reliability for the verbal booklets of this study was calculated by STS for the verbal A booklets as: fluency = .95, flexibility = .86, and originality = .86. The inter-rater reliability for the figural booklets of this study was established by calculating reliability coefficients of scoring for the streamlined figural tests by this investigater with scores on a random sample of 20 Figural B booklets by three other independent scorers. The first comparison for inter-rater reliability of standard scores for this rater and an experienced rater at STS scoring service, yielded coefficients for fluency, r = .99, and originality, r = .90, that were comparable to values reported by Torrance and Ball (1984). However, the values for titles, elaboration, and closure were all less than or equal to .78 (see Table F-7, Appendix F).

The booklets were, therefore, scored a third time by a graduate student at the University of Georgia. These inter-rater reliability coefficents were (F=Fishkin, G=Georgia, S=Scoring service): fluency $r_{F/G} = .99$, $r_{G/S} = .98$; originality $r_{F/G} = .67$, $r_{G/S} = .71$; titles $r_{F/G} = .91$, $r_{G/S} = .58$, elaboration $r_{F/G} = .89$, $r_{G/S} = .49$, closure $r_{F/G} = .96$, $r_{G/S} = .61$, and total of creative strengths $r_{F/G} = .81$, $r_{G/S} = .55$ (see Table F-7, Appendix F).

The discrepancy in scoring between this author and the first STS scorer (see Table F-7, Appendix F) is in part explained by differences in usage of the manual. Experienced STS scorers were using current updates of the scoring procedures. For example, an updated interpretation of resistance to premature closure of the drawings included the term "conceptually closed" (T. Safter, personal communication, December 5, 1987). The STS scorer estimated elaboration points, whereas this author, as an inexperienced scorer, counted the points for elaboration. Inexperienced scorers tend to score more generously for elaboration than experienced scorers (O. F. Anderhalter, personal communication, June 15, 1988). Moreover, streamlined scoring for figural B was more susceptible to scorer interpretation errors than for figural A because the explicit guideline examples were available in a workbook published only for figural A (Ball & Torrance, 1984).

The STS indicated that it would be helpful to reassess the sample of 20 Figural B booklets because the streamlined manual (Torrance & Ball, 1984) was under revision. The comparison was considered to be helpful to STS to highlight areas of the manual that were in need of clarification (O. F. Anderhalter, March 24, 1988, June 15, 1988; J. D. Kauffman, January 8, 1988; personal communications). For the reassessment, the set of booklets were scored by the STS senior scorer who revised her usual scoring standards based on current updates available to STS. The senior scorer from STS followed a literal interpretation of the manual (Torrance & Ball, 1984), and counted elaboration points in order to replicate the procedures used by this author. These reliability coefficients for the third comparison, unlike the preceding sets of correlations, were calculated on raw score data rather than standard score data (see discussion below). These raw score coefficients were: fluency = .99, originality = .96, titles = .82, elaboration = .76, closure = .89, and total of creative strengths = .66 (see Table F-8, Appendix F). These resultant correlations indicated that this author interpreted the manual adequately and obtained acceptable inter-rater reliability, all $r \geq .76$ (see Table F-8, Appendix F).

It is possible that these inter-rater reliability coefficients may reflect a somewhat higher level of reliability for the verbal tests of this study than could perhaps be present in the set of data which was used for the final analyses. That is, these inter-rater reliability calculations were conducted on the full range of scores in the set of data before the ceiling of 160 was imposed. The variability was restricted in the range of scores, therefore resulting in a set of the data which could perhaps have had a lower inter-rater reliability had it been analyzed.

The obtained inter-rater reliability for this study is at a level that is acceptable for research purposes for all the measures. It should be noted, however, that the values of the final reliability coefficients and of scoring for elaboration from this small inter-rater reliability study are markedly different from those reported in the streamlined manual (Torrance & Ball, 1984). The reliability coefficients from five studies at five different grade levels yielded coefficients $\geq$ .90 for all five norm-referenced scoring categories with the exception of one at .78, closure, from a sample of college students (Torrance & Ball).

The level of scores for elaboration based on streamlined scoring bears comment. The manual (Torrance & Ball, 1984) reported that a set of booklets was scored for fluency, originality, and elaboration by the same scorer under standard (Torrance, 1974a) procedures and six months later by streamlined scoring procedures (1984). The inter-rater reliability coefficients were .92, .94, and .92 respectively. Use of the streamlined procedures of estimating elaboration points rather than counting points resulted in a significantly higher number of points. The reverse occurred in the present study: a significantly lower score was obtained when elaboration points were estimated rather than counted. Rosenthal et al. (1983) did not include elaboration measures in their inter-rater reliability study. There is not a prior study for comparison of the effects of different rater judgments of elaboration in a sample of gifted children. Halpin and Halpin (1974), however, found the greatest discrepancy of the seven TTCT measures between their scorers on the figural elaboration scores when scoring booklets of university level subjects. These results, then, suggest that differences between raters appear to be most marked on elaboration scores.

Effects on Reliability Due to
Conversion to Standard Scores

"The most powerful scores, statistically, are standard scores derived from the properties of the normal probability curve and preserving the absolute differences between scores" (Isaac & Michael, 1981, p. 105). However, the standard scores for the TTCT are not a true standard score conversion based upon deviation from the mean by use of scores (Isaac & Michael). The TTCT "standard scores," rather, are conversions "determined on the basis of grade level national norms" (Torrance & Ball, 1984, p. 16) and, therefore, are more closely related to grade level norms than to conversions to a standard score. The conversion to "standard scores" moreover for the titles and the closure measures introduces a source of unreliability because of the limited spread in those scores. That is, the conversion introduces an additional source of unreliability by accentuating differences between scores when the raw scores are actually close together in value. Therefore, use of raw scores is preferred in calculating reliability coefficients for the TTCT (O. F. Anderhalter, personal communication, June 15, 1988, see Appendix A). The inter-rater reliability coefficients for the first three comparisons were recomputed with raw score data (see Table F-6). The results of these calculations confirmed Dr. Anderhalter's observation that use of standard scores changes the inter-rater reliability coefficients. A comparison of the correlations of raw score data presented in Table F-6 and those of standard score data presented in Table F-7 provides clear evidence for the advantages of raw scores for reliability calculations of TTCT data. The comparison shows that the raw score analysis revealed a stronger relationships for 11 of the 12 correlations which changed. The distortion due to standard score conversion is most clearly shown in the coefficients for closure: this author and STS, standard score r = .47, raw score r = .63; STS and Georgia, standard score r = .49, raw score r = .62.

Standard scores, because they preserve the integrity of differences between scores are appropriate in any interval scale statistical calculations. Such true standard scores offer the advantage for the user to make clear comparisons among a variety of different test scores "if the reference groups are equivalent. {However, a major disadvantage is that if the} data are markedly skewed, they are inappropriate . . ." (Isaac & Michael, 1981, p. 105). In retrospect, although true standard scores, when available, are the preferred unit for statistical analysis, the TTCT raw scores should be preferred to the "standard scores." In this study, performed on a sample of gifted children, the creativity data were found to be markedly skewed (see Results and Discussion, Chapters IV and V, pp. 141-145, 209-211), and therefore the use of the standard scores was found to be very inappropriate. However, because the TTCT standard scores are not true standard scores, but rather are grade level standard scores, it is not clear if this restriction applies. These findings, then, suggest that the raw scores may be more appropriate than the TTCT standard scores for research studies because of the greater reliability of the raw scores, because of the grade level nature of the conversions to the standard scores, and because of the abnormal distribution of scores when many in the sample of scores are highly creative. The last concern is likely to be a problem in many studies using the TTCT which would often expect to find disproportionate numbers distributed in the upper ranges of the scores. However, the recommendations from Torrance (1974a; 1974b) and especially from Torrance and Ball (1984) are too brief to explain the rationale for which scores might be preferred for various purposes.

This author concurs with Torrance (1974a) and Torrance and Ball (1984) that the standard or earlier T-score (1974a; 1974b) conversion is recommended when the purpose is to determine the relative strengths of the different kinds of creative abilities of a given individual or when the purpose is to assess an individual's performance against a comparison group. The standard scores would, therefore, be recommended if the purpose is primarily for identification of creative individuals. In 1974, Torrance recommended against use of composite total verbal and/or figural scores even though "reliabilities are generally higher for such total scores than for the separate . . . scores" (Torrance, 1974b, p. 56). The new Creativity Index, which is a composite score, however, is recommended by Torrance and Ball (1984). The results of this study show it to be more reliable than the separate scores. The Creativity Index, by utilizing the creative strengths, offers a meaningful composite TTCT score comparison which should be a helpful predictor of creative behaviors.

In summary, lower alternate forms reliability coefficients were found in this study for verbal and figural TTCT than for those reported for the original scoring (Torrance, 1974b). Lower values were also found for inter-rater reliability comparisons. These findings, moreover, provide evidence to confirm the recommendation that it is important to use a single rater for consistency of the level of scoring in a single study or for purposes of identification (Halpin & Halpin, 1974; Rosenthal, et al., 1983).

However, there are several factors present which lead to the conclusion that the obtained reliability coefficients in this study are underestimates, particularly because the range is restricted to gifted children. Moreover, when a test is able to predict a relevant criterion to a satisfactory level, then "low reliability is not a serious problem (e.g., tests of creativity)" (Isaac & Michael, 1981, p. 123). Because the criterion validity of the TTCT is adequate, the low reliability coefficients, with the possible exception discussed above of closure, should not be a problem, unless the results are used for purposes of identification where a fixed cut-off score will serve to exclude children (see discussion below).

Identification of the Creatively Talented

For identification purposes these results offer support to Torrance and Ball's (1984) suggestion that the overall Creativity Index should be used. The Creativity Index is calculated by adding the raw score total of creative strength indicators to the averaged standard scores for the five other measures from the figural battery, the five norm-referenced scores. The Creativity Index, is, as would be expected, more reliable, and it is also more valid than any of its component subscales alone because it reflects the important contributions of the creative strengths.

For identification purposes, then, the Creativity Index provides a single score of promising reliability and validity. Evidence is presented by the results of this study which supports the recommendation by Chase (1985) that a single total or composite score would suffice for the verbal measures (see Chapter V, implications for the verbal measures). Use of a single rater, or, at the very least, rigorous control of inter-rater reliability, and local norms have been recommended (Halpin & Halpin, 1974; Rosenthal, et al., 1983; Torrance, 1987). The TTCT is a popularly used measure which yields strong evidence of predictive validity in research studies when scores of groups of individuals are compared (Howieson, 1981; Torrance, 1972b; Torrance & Ball, 1984). The level of reliable assessment of an individual's creative potential, however, is, in the TTCT, as in any measure of creative production subject to considerable variation. Because of problems in reliability including variability in scores in relation to effects of motivation identified by prior research (Elkind, et al., 1970; Halpin & Halpin, 1974; Rosenthal, et al., 1983) and by the results of this study, this author strongly agrees with Davis and Rimm (1985).

> Central to the appropriate use of divergent thinking tests such as the Torrance Tests or personality/biographical inventories is the recognition that (1) scores from a single creativity test should be combined with other information, such as teacher ratings of creativity or scores on a second creativity test, in order to reach a valid decision, and (2) low creativity test scores absolutely must never be used to eliminate children from G/T programs. Creativity tests are not perfect; there simply are too many types of creativity and creative people. However, creativity tests can identify creatively gifted children, majority and minority, who may not be identified in other ways (p. 265).

Contributions of this Study to Validity Information

There are three sources of data available from this study which bear upon construct validation of the new streamlined scoring of the figural TTCT: 1) information about the relationship of the ten creativity variables with each other from the intercorrelation tables (Tables P-1 and P-25, Appendix P), 2) information about the strength of the relationship of the various variables as predictor variables from the multiple regression equations (Tables P-8, P-17, P-26, and P-27, Appendix P), and 3) information about the effects of the team creative problem solving (CPS) treatment variable, Odyssey of the Mind (OM), on the TTCT dependent variables (see results and discussion, Chapters IV and V).

The streamlined scoring procedures are relatively new and "thus far only a few predictive, concurrent and construct validity studies have been conducted and reported" (Torrance & Ball, 1984). A second measure of cognitive creativity, Similes, therefore was included in this study because the validity information on the new TTCT scoring was still sparse. The inclusion of Similes served two purposes: it provided a creative measure independent of the TTCT batteries to provide information about possible treatment effects in the study and it served to provide a measure to obtain concurrent criterion validity comparison for the various TTCT variables. Criterion validity information, then, is available from comparisons of Similes to the various TTCT variables and also from comparisons between the streamlined figural battery measures and those of the verbal battery.

Construct Validity

The intercorrelation of the ten creative dependent variables (Table P-1, Appendix P) and the multiple regression analyses for these variables (Tables P-8, P-17, and P-27) yield evidence of construct validity of the new scoring variables of abstractness of titles (titles), resistance to premature closure (closure), and total indicators of creative strengths (creative strengths). Titles carries significant correlation with the creative strengths, the most important predictor variable for titles found from the multiple regression analysis. It was, however, surprising that Similes did not carry significant relationship with abstractness of titles. The scoring criteria for Similes and abstractness of titles appear on first glance to be comparable. The highest score for a title is a "3," described as "abstract but appropriate . . . capturing the essence of the picture, going beyond what is seen . . ." (Torrance & Ball, 1984, p. 19). The highest score for a simile is a "5," described as "Exceptional quality. Very clear or unusual idea. A unique response that has an arresting effect on the reader and results in a desire to savor" (Schaefer, 1971, p. 15). Torrance (1979) considered ability in abstractness of titles to reflect the person's ability to highlight the essence of what was drawn. In an unpublished study by Torrance, reported in Torrance and Ball (1984), however, titles, as well as fluency and originality, carried significant correlation with Similes. On closer consideration of the scoring criteria for titles and for Similes, however, differences between these emerge. A high score in abstractness of titles reflects the ability to synthesize, to express the essential quality. It may not necessarily, like a high score on Similes, be emotionally expressive, which is, in fact, a different TTCT scoring category under the creative strengths.

The intercorrelation data for closure showed it to relate with Similes, figural and verbal originality, figural fluency, elaboration, and the total of creative strengths. The regression, however, did not show figural originality as a significant predictor of resistance to closure. Because differences in grade as a supressor variable accounted for a major portion of the variance of closure, the regression equation yielded figural fluency as the only predictor of the TTCT variables (see Table P-8, Appendix P). Torrance (1979) described resistance to premature closure as a measure of the creative quality of psychological openness. Psychological openness is considered by Torrance to be closely related to deferring of judgment, a quality that is critical to the CPS process as discussed in Chapters I-III. If so, then, an increase in closure scores would be expected to be associated with CPS training. The regression equation confirms this prediction: years of prior experience on an OM team was a significant predictor of closure scores. Table P-27 shows the regression analysis where the six control variables were omitted in order to evaluate the contributions of the 20 creative predictor variables more clearly. In this analysis, without the contributions of grade and OM as suppressor variables, elaboration, Similes, and to a lesser extent the creative strengths emerged as contributors as well as figural fluency to the predictor equation for closure.

The creative strengths represented the raw score total of the 13 indicators of creative strength: emotional expressiveness, storytelling articulateness, movement or action, expressiveness of titles, synthesis of incomplete figures, synthesis of lines or of circles (activity 3), unusual visualization, internal visualization, extending or breaking boundaries, humor, richness of imagery, colorfulness of imagery, and fantasy. Credit for any of these indicators is assigned when the specific criteria of the scoring cateogry are met. These points for the creative strength indicators should be awarded sparingly, with close attention to the criteria, because Torrance and Ball (1984) have given the strengths importance in the calculation of the Creativity Index (O. F. Anderhalter, personal communication, June 15, 1988). The creative strengths had a significant correlation with Similes and with all of the variables from the figural battery. There was a moderately strong correlation between the creative strengths and elaboration, r=.53. The stepwise multiple regression equation in Table P-8, however, did not load figural originality or closure as significant predictors but did show pretest variables in total of creative strengths, posttest variables for titles and elaboration, and grade, and sex, to be significant predictors of the posttest scores for the creative strengths. In the analysis of the creativity predictor variables (Table P-27), figural originality was shown to carry some weight in the regression and verbal originality was a suppressor variables.

The intercorrelation data of the figural Creativity Index with the other two measures of the abbreviated creative construct, Similes and the mean of the verbal measures (Table P-12), show a significant correlation of the Creativity Index with Similes but not with the verbal TTCT measures. The regression equation for the posttest Creativity Index (see Table P-17) shows confirmation of the strong relation to Similes; the pretest variables for Similes, figural fluency, and elaboration were strong predictors of the Creativity Index. The mean verbal TTCT equation showed pretest verbal fluency to have the greatest contribution with some contribution from posttest Similes and from the control variables of grade and teacher.

The data from the intercorrelation and multiple regression analyses, then, show that the new variables in the streamlined scoring--titles, closure, and creative strengths, and the cumulative Creativity Index--in general bear relationship with Similes but not with the verbal TTCT divergent production measures of fluency, flexibility, and originality. These findings show that the new measures add dimension to the TTCT figural battery that take it beyond a measure of divergent production thinking to a measure that taps other important aspects of creativity. This author considers the scores obtained from Similes to reflect the richness of verbal creative behavior to a greater extent than the scores of the verbal TTCT battery. The divergent thinking categories of fluency, flexibility, and originality, scored as a statistically infrequent response, are too limited to reflect the essence of creative behaviors shown by Similes and by the new streamlined figural battery. These impressions find further confirmation in results of the treatment effects of this study.

Criterion Validity

In this study, the children involved in the enrichment program had varying amounts of prior experience with creativity training. Some were starting their first year of the program and others were starting their fourth or fifth year of participation. Moreover, the children had varying amounts of prior experience with the Odyssey of the Mind (OM) program. The OM program is a team creative competition program which uses creative problem solving (CPS) techniques. It is described in Chapter III and Appendix C. The effects of these experience factors were tested by two 3x2x2 mixed model MANOVAs, one for Experience in the Enrichment Program x OM x Time, and the other for Prior Experience in OM x OM x Time. The analyses showed these experience factors to yield significant differences in pretest performance. Both experience factors, moreover, had the same pattern of variables reflecting the prior training conditions in creativity. These were an increase in scores on the creative strengths, Similes, closure, figural elaboration, and figural originality, but a decrease in verbal flexibility for children who had greater experience in creativity training. Moreover, significant differences in pretest scores were present between the children electing to join an OM team and those who did not. The children who joined a team in January were, on the average, higher in Similes, closure, and total creative strengths than children who did not participate in OM (see Chapter IV). These results indicated that, despite the low reliability shown for closure, it bears a valid relationship to the criteria of experience in creativity training. That is, closure appeared in a cluster of creative abilities including Similes and the creative strengths which differentiated children who chose to join a creative competition team from those who did not. This cluster of abilities, along with figural elaboration and figural originality, was associated with differentiating between children who were experienced with previous creativity training and those who were not (see Chapter IV and V).

These pre-test differences in experience and in OM effort groups were shown in analyses of the abbreviated creativity construct using Similes, the figural Creativity Index, and the mean scores of the three verbal measures. Similes and the figural Creativity Index were responsive to the effects of prior experience in creativity training and pretest differences in the OM groups. The data from this study provide evidence of the usefulness of the Creativity Index. The Index showed higher reliability than any of its component scores alone and, as would be expected, showed responsiveness to the effects of creativity training.

In examining the effects of time, there was a significant main effect for time across all subjects. Scores on the figural Creativity Index and on Similes increased over time, whereas with the exception of the OM high effort group, the verbal measures decreased. In examining the individual measures for their effect from pretest to posttest, there was a marked increase in figural originality and Similes, an increase in figural fluency, and decrease in scores on figural elaboration, verbal fluency, and verbal originality. It appears that the verbal measures and figural elaboration were susceptible to the context effect (see above) as a likely inhibitor of motivation for the expenditure of creative energy on these tests. Similes and the figural battery were administered before the verbal tasks and, with the possible exception of elaboration, did not seem to suffer the same negative effects.

The effect of the OM treatment was shown in the verbal test scores, particularly in verbal flexibility and in Similes, but not on the figural battery or Creativity Index analyses. Despite the potent context effect of depression of posttest creativity scores, there was a significant difference between the OM effort groups. That is, the group of children who were high in effort on their OM teams did not show the decline in posttest verbal scores and were significantly higher in Similes than the group who showed low effort in their work in OM (see Chapters IV and V). The majority of the children in OM in that year were working on problems that were more of a verbal nature than of a non-linguistic focus. Because their OM experiences and training in creativity focused more on verbal creativity than on expression in a figural mode, it is understandable that the verbal scores and Similes reflected the OM training where the figural scores did not. Cohen (1987) in a case study of five OM teams and their coaches found that the team members and coaches were shown to be more creatively talented by the verbal TTCT measures than by the figural TTCT. Three of the five teams worked on problems that were geared more toward the dramatic arts, but all five teams were stronger in the verbal than the figural TTCT. These findings are consistent with results reported by Rose and Lin (1984) where the TTCT scoring category most affected by CPS was verbal originality, while figural scores were most strongly affected by the Purdue Creative Thinking Program (Feldhusen, Speedie & Treffinger, 1971) and other school creativity training programs.

Summary and Conclusions

In conclusion, then, these data offer an expanded base of reliability and construct and criterion validity information for the new streamlined measures. In this appendix, the reliability and validity base for the TTCT derived from prior studies is briefly discussed, and is focused more closely on the new streamlined procedures. The sensitivity of the tests to the effects of warm-up and other motivational influences is mentioned. The detrimental effect on TTCT scores found by Elkind et al. (1970) of interrupting ongoing activities that are intrinsically interesting and important to the children is discussed as a likely effect on this study. In attempting to establish acceptable inter-rater reliability of the figural measures for this study, comparisons between four different raters were produced. The results of these reliability comparisons are discussed. The possible effects of using the raw scores in preference to the standard scores for reliability comparisons and for research studies involving gifted children is discussed.

These findings suggest a need for additional clarity and guidance in the TTCT manuals for recommended uses for standard versus raw scores. It might be helpful if some discussion in the manual could be directed to uses for research as opposed to uses for indentification of individuals. These findings suggest that the raw scores may be the more appropriate unit for research purposes where the study is exploring the nature of creative thinking. However, the Creativity Index is a stronger and more reliable measure than any of the component scores. Use of composite scores for the verbal and for the figural tests enabled the treatment effects in this study to emerge because the variability contributed by scores of the verbal measures that were highly incorrelated was thereby controlled. The findings support the recommendations by Torrance and Ball (1984) that the Creativity Index should be the appropriate figural measure for identification purposes. However, because of problems of variability attributed to motivation, administration, and scoring inherent in any open-ended test of creative production, this author concurs with Davis and Rimm (1985) that creativity test scores must be used with another type of measure such as teacher ratings or self-report inventories of aspects of personality associated with creativeness. Moreover, as Halpin and Halpin (1974) recommended, if local raters are used, because the levels of the scores may be significantly different among raters, local norms should be generated for purposes of comparison of individuals.

Finally, this study offers contributions to the base of validity information for the new streamlined variables of abstractness of titles and resistance to premature closure. The addition of the creative strengths and resultant Creativity Index offer measures that reveal important information about the children's creative functioning and are sensitive to the effects of prior training in creativity.

APPENDIX F

DISTRIBUTION OF SUBJECTS WITH

RESPECT TO EFFORT IN OM

Table F-1

Distribution of Subjects with Respect to Odyssey of the Mind (OM) Effort and Experience in the Enrichment Program

	OM Effort			
Experience in Program	Non-OM	OM-lo	OM-hi	Total
Less than 13 months				
Creativity Data	18	11	11	40
Affective Data	17	9	11	37
13 months or more				
Creativity Data	33	21	22	76
Affective Data	32	20	22	74
Total				
Creativity Data	51	32	33	116
Affective Data	49	29	33	111

Table F-2

Distribution of Subjects with Respect to Odyssey of the Mind (OM) Effort and Prior Experience in OM

	OM Effort			
Experience in OM	Non-OM	OM-lo	OM-hi	Total
No Prior OM				
Creativity Data	39	12	10	61
Affective Data	37	11	10	58
Prior OM				
Creativity Data	12	20	23	55
Affective Data	12	18	23	53
Total				
Creativity Data	51	32	33	116
Affective Data	49	29	33	111

Table F-3

Distribution of Subjects with Respect to Odyssey of the Mind (OM) Effort, Grade Level, and Sex

		OM Effort						
		Non-OM		OM-lo		OM-hi		Total
Grade		Girls	Boys	Girls	Boys	Girls	Boys	
2	Creativity Data	7	4	-	1	1	1	14
	Affective Data	7	4	-	2	1	1	15
3	Creativity Data	8	8	1	6	-	5	28
	Affective Data	7	8	1	6	-	5	27
4	Creativity Data	6	9	5	11	4	10	45
	Affective Data	6	8	4	8	4	10	40
5	Creativity Data	5	4	4	4	4	8	29
	Affective Data	5	4	4	4	4	8	29
Total	Creativity Data	26	25	10	22	9	24	116
	Affective Data	25	24	9	20	9	24	111

Table F-4

Distribution of Subjects with Respect to Odyssey of the Mind (OM) Effort, Grade, and School* in Which Their Enrichment Classes Were Located

		OM Effort												
		Non-OM				OM-lo				OM-hi				Total
		Grade				Grade				Grade				
School		2	3	4	5	2	3	4	5	2	3	4	5	
1	Creativity Data	1	2	5	1	1		2		1		3	4	20
	Affective Data	1	2	4	1	1		2		1		3	4	19
2	Creativity Data	1						1	1					3
	Affective Data	1							1					2
3	Creativity Data	5	3	3	2	1	3	4	3	1	2	9	4	40
	Affective Data	5	3	3	2	2	3	4	3	1	2	9	4	41
4	Creativity Data			1	1		2	3	1		3	1	4	16
	Affective Data			1	1		2	3	1		3	1	4	16
5	Creativity Data	2		1	1		1	1	1					7
	Affective Data	2		1	1		1		1					6
6	Creativity Data		4	1	1									6
	Affective Data		3	1	1									5
7	Creativity Data	1			1			3	1			1		7
	Affective Data	1			1			2	1			1		6
8	Creativity Data			3				1						4
	Affective Data			3				1						4
9	Creativity Data		3	2										5
	Affective Data		3	2										5
10	Creativity Data	1	4	1				1	1					8
	Affective Data	1	4	1					1					7
Total	Creativity Data	11	16	17	7	2	6	16	8	2	5	14	12	116
	Affective Data	11	15	16	7	3	6	12	8	2	5	14	12	111

*Enrichment classes were located in:
Schools 1 - 2 were taught by Teacher 1 - Anne Fishkin
Schools 3 - 4 were taught by Teacher 2 - Kathleen Kaufman
Schools 5 - 10 were taught by Teacher 3 - Barbara Sharp

Table F-5

Distribution of Subjects with Respect to Odyssey of the Mind (OM) Effort and Parental Permission to be in the Study

	OM Effort			
Permission	Non-OM	OM-lo	OM-hi	Total
Permission Granted				
Complete Data Obtained from Either Creative or Affective Battery				117
Creativity Data	51	32	33	116
Affective Data	49	29	33	111
Permission Granted-Not Tested			1*	1
Student Moved	5	1		6
Student Withdrew from Enrichment Program	7 + 1**	1		9
Total				133
Permission Not Granted				
Remained in Program	5		2	7
Student Moved	1			1
Student Withdrew from Enrichment Program	2			2
Total				10

*Parental permission was granted; student refused to take tests.
**Permission granted for creative but not affective tests.

Table F- 6

Inter-Rater Reliability of Three Raters: Streamlined Scoring of Figural TTCT,

Random Sample of 20 Cases-Comparisons of Raw Scores

Rater[a]	Figural Dependent Variables											
	Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
$\bar{X}$ and SD of Raters												
Author	22.7	6.6	22.7	5.1	7.5	3.3	12.4	2.9	13.0	3.2	9.0	3.4
Scholastic Testing Service	22.8	6.4	22.8	5.3	6.1	1.9	7.6	1.5	12.9	2.9	8.0	3.1
Georgia	22.6	6.5	20.4	5.1	6.5	3.5	13.5	2.5	13.3	3.1	8.9	2.7
Inter-Rater Correlations												
$r_{Author/STS}$	.99		.96		.82		.63		.89		.66	
Author/Georgia	.99		.76		.91		.91		.93		.81	
$r_{STS/Georgia}$	.99		.76		.62		.62		.86		.55	

[a]Raters-Author, A. Fishkin; Scholastic Testing Service, a trained professional scorer under contract with Scholastic Testing Service; Georgia, a graduate student at University of Georgia referred to the author by Torrance Center for Creative Studies, Athens, Georgia.

Table F-7

Inter-Rater Reliability of Three Raters: Streamlined Scoring of Figural TTCT,

Random Sample of 20 Cases-Comparisons of Standard Scores

	Figural Dependent Variables											
	Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
Rater[a]	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
$\bar{X}$ and SD of Raters												
Author	108.8	17.4	145.0	19.3	111.4	18.6	125.3	17.9	116.3	17.7	9.0	3.4
Scholastic Testing Service	107.9	17.1	151.1	22.1	104.8	13.4	89.4	11.3	103.0	12.5	8.0	3.1
Georgia	108.9	16.6	185.8	19.1	105.5	19.6	132.5	16.6	117.2	18.3	8.9	2.7
Inter-Rater Correlations												
$r_{Author/STS}$	.99		.90		.78		.47		.66		.66	
Author/Georgia	.99		.67		.91		.89		.96		.81	
$r_{STS/Georgia}$	.98		.71		.58		.49		.61		.55	

[a]Raters-Author, A. Fishkin; Scholastic Testing Service, a trained professional scorer under contract with Scholastic Testing Service; Georgia, a graduate student at University of Georgia referred to the author by Torrance Center for Creative Studies, Athens, Georgia.

Table F-8

Inter-Rater Reliability of Two Raters: Streamlined Scoring of Figural TTCT, Random Sample of 20 Cases-Comparisons of Raw Scores[a]

	Figural Dependent Variables											
	Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
Rater[b]	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
$\bar{X}$ and SD of Raters												
Author	22.7	6.4	22.7	5.0	7.5	3.2	12.4	2.8	13.4	2.8	9.0	3.4
Scholastic Testing Service	22.9	6.3	22.8	5.2	7.5	3.1	10.9	2.3	12.9	2.9	8.0	3.0
Inter-Rater Correlations												
Author/ Scholastic Testing Service	.99		.96		.82		.76		.89		.66	

[a]Source-Dr. O. F. Anderhalter, Director, Research and Development Division, Scholastic Testing Service.

[b]Rater-Author, A. Fishkin; Scholastic Testing Service, S. Lillard, Scholastic Testing Service Senior Scorer.

Table F-9

Test-Retest Reliability (Stability) of Instruments[a]

Instrument	Pretest $\bar{X}$	Pretest SD	Posttest $\bar{X}$	Posttest SD	$\underline{n}$ for r	Pretest/ Posttest r^b
Creative Instruments						
Similes	43.43	16.58	51.74	17.43	118	.53
TTCT Verbal						
Fluency	125.60	21.34	120.42	19.38	117	.49
Flexibility	133.87	17.82	133.15	19.06	117	.51
Originality	126.71	20.05	110.16	16.68	117	.45
Mean Verbal	128.73	19.11	121.24	17.43	117	.53
TTCT Figural						
Fluency	101.92	16.80	108.79	16.15	117	.37
Originality	119.70	19.78	137.83	20.54	117	.45
Titles	110.65	17.65	113.47	18.31	117	.32
Elaboration	129.26	19.55	118.60	20.64	117	.42
Closure	108.71	11.75	111.31	16.79	117	.18
Creative Strengths	9.16	2.95	9.73	3.14	117	.40
Creativity Index	123.21	12.23	127.73	11.55	117	.45
Affective Instruments						
Sears	182.17	30.67	179.54	33.22	114	.69
I+	13.59	2.30	13.87	2.50	114	.52
I-	9.73	3.31	10.49	2.95	114	.53
Ideas	34.17	6.77	36.18	6.22	116	.44

Note. Standard scores are reported for all TTCT measures except for the creative strengths and Creativity Index. The total of creative strengths, Similes, and the affective measures are reported as raw scores. The Creativity Index is calculated by adding the raw score total of creative strengths to the average of the standard scores for the five norm-referenced scores.

Note. These reliability estimates are probably underestimates of the test-retest reliability for these instruments in that the time interval was five months between testing sessions with intervening treatment effects for the different groups in the study.

[a]All TTCT standard scores at $\leq$ 160.
[b]All Pearson correlation coefficients are significant at $p < .01$ with the exception of Closure.

Table F-10

Descriptive Statistics of Distribution Ranges and Skewness for the Creative and Affective Instruments (Pre- and Posttest)[a]

	Descriptive Statistics				
Measure	Mean	Std. Dev.	Minimum	Maximum	Skewness
Creative Instruments					
Similes	43.43[b] 51.74	16.58 17.43	8 6	92 91	.34 -.17
TTCT Verbal					
Fluency	128.69 120.65	28.47 19.90	84 69	232 172	1.24 .22
Flexibility	137.25 134.41	24.78 21.21	93 63	228 183	1.15 -.02
Originality	129.88 110.29	27.13 17.11	92 82	231 175	1.36 .93
TTCT Figural					
Fluency	101.92 108.79	16.80 16.15	67 75	142 157	.41 .34
Originality	119.89 141.45	20.20 25.90	75 76	170 221	.29 .14
Titles	110.65 113.57	17.65 18.58	65 78	160 166	.17 .72
Elaboration	129.81 118.60	20.51 20.64	72 72	170 160	-.08 -.15
Closure	108.71 111.31	11.75 16.79	73 77	130 155	-.49 .37
Creative Strengths	9.16 9.73	2.95 3.14	3 2	16 18	.21 .10
Affective Instruments					
Sears	182.04 176.54	30.82 33.22	90 65	234 232	-.63 -.85
I+	13.59 13.87	2.30 2.50	7 5	17 17	-.57 -1.09
I-	9.73 10.49	3.31 2.95	3 3	17 17	-.04 -.35
Ideas	34.17 36.18	6.77 6.22	15 20	50 50	-.28 -.15

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes, the total of creative strengths, and the affective measures are expressed in raw scores. See Chapter III, Instrumentation.

[a]For these analyses, scores for the 17 students whose standard scores were greater than 160 are treated as obtained.
[b]Pretest statistics are above posttest data.

APPENDIX G

LETTERS SEEKING PARENTAL PERMISSION

Edmond Public Schools
Special Services Center
215 N. Boulevard
Edmond, Oklahoma 73034

September 4, 1985

Dear Parents:

Within the area of gifted education the enhancement of creative potential of the student is a goal. Mrs. Kaufman, Mrs. Sharp, and I wish to ascertain the effect of activities offered in the gifted classes on the creativity level demonstrated by the children.

In order to measure the effect and changes in creativity those children for whom parents grant permission will be given a variety of creativity and affective surveys. All tasks will be presented in a relaxed, matter-of-fact manner so that the testing process would be a valid measure and agreeable to the children. Children tend to enjoy doing the activities in the creativity test battery. Pretests will be administered in three sessions during part of enrichment classes in the fall and posttests will again be administered in the spring for a total of 6½ hours of testing during the school year.

The aspect of increasing levels of creativity through activities is a special interest of mine. I believe it to be worthy of study and feel very knowledgeable in this area. For this reason I have chosen it as a topic for study at Oklahoma State University. In my study I would like to include the results of these tests. No student names would be mentioned. I would be, however, available to discuss the meaning and validity of an individual child's test results with the child's parent(s).

We are also asking your permission at this time to administer enrichment program evaluation measures which had routinely been administered in past years. Mrs. Kaufman, Mrs. Sharp or I would be glad to share data from previous years' evaluations with you upon request.

If you are willing for your child to participate in testing sessions during enrichment class time, please sign the form below and return to me by September 20th in the enclosed, stamped envelope. If you have any questions about the proposed group-administered measures to be given to the children, please leave a message at this office, 341-0405, and I will return your call, or feel free to call me at home, 755-6387. Thank you for your help.

Sincerely yours,

Anne Fishkin

Anne Fishkin
Enrichment Teacher

My child ________________________________ may participate in enrichment program testing sessions during the 1985-1986 school year in the following manner:

/_/ All creativity and affective tests, pre- and posttests and regular Lab end-of-year surveys. Total testing time: 7½ hours in 8 sessions during enrichment classes.

/_/ Permission to use results in study with student confidentiality assured.

/_/ Creativity tests and regular end-of-year surveys only. Time: 5½ hours.

/_/ Regular Lab end-of-year surveys only. Time: 1 hour in 2 sessions.

/_/ Please arrange for a feedback session with me for my child's test results at a later date.

/_/ I do not want my child to take any of the above-listed evaluations.

________________________ Date

________________________ Parent's Signature

Enrichment Program
Evaluation Measures for 1985-1986 School Year

Rationale for the Study

Growth in enrichment program students is frequently observed by program teachers in critical thinking, creativity, and affective (feeling) behaviors such as in self-confidence, self-assertiveness, willingness to take risks, leadership, and positive attitudes toward school and learning. End-of-the-year evaluations from students, coaches, parents and other teachers tend to support these observations of growth. However, these are only indirect indicators of growth in affective areas. A review of the recent literature in the area of creative and affective changes in gifted children reveals a surprisingly small number of studies and a need to use more reliable and precise ways of measuring changes which might be associated with creativity training in young people.

Testing Procedures

Testing procedures for the research study will take approximately 1 hour out of each of 6 Lab classes during the school year. Only students for whom we have received written parental permission may participate in the testing program. During testing sessions those students for whom we do not have written permission will be working on appropriate enrichment materials for credit, such as their choice of various learning centers or working in a critical thinking workbook.

The measures of creativity will be the Torrance Tests of Creative Thinking, figural and verbal, and a brief measure of literary talent, Similes. These measures will require 2 hours of administration time in two one-hour sessions given in mid-October and again in late April for a total of 4 hours. The Torrance measures are presented in a game-like class environment and are usually motivating and interesting to take. The affective measures will require another 60-85 minutes to be given in late October or early November and again in early May. These measures will consist of 5 surveys related to how the student feels about the self as a creative person, attitude toward the self as an effective person and toward leadership qualities. In all events, an individual student's name and test results would remain confidential.

Use of the Results

Analysis of the data will be made after the end of the school year so the tests cannot be returned to you. If you desire feedback about your child's performance on these research instruments please indicate so on the attached permission slip. The test results will be scored, analyzed and discussed for the study. If meaningful results are obtained they would probably be shared in an article and presentations in a professional journal or at a conference. If desired, the results could be shared with Edmond parents and/or faculty.

Program Evaluation Measures

A regular enrichment program procedure for the past six years was to administer some program evaluation measures to all the children after OM creative problem solving work and at the end of the year. These measures, evaluating participation in the program, and evaluation of the teacher, usually took a total of 45 minutes. The additional tests we are proposing this year can give us greater insight in improving the gifted program. In general, the children enjoy the opportunity to express their ideas and feelings about the program. We assure them that their opinions are important, the surveys are read, and from their ideas and the evaluations of their parents we continue to look at ways to improve the program. If requested, Mrs. Kaufman, Mrs. Sharp or Mrs. Fishkin would be available to share analyses of data from previous years' program evaluations.

Edmond Public Schools

Special Services Center
215 N. Boulevard
Edmond, Oklahoma 73034

October 29, 1985

Dear

Thank you for giving permission for your child to be included in the special enrichment program study this school year. As we discussed by telephone, the written, verbal portion of the Torrance Tests of Creative Thinking will yield more valid results and interpretation if administered on an individual basis to those children who are in 3rd grade or younger. Thank you for giving permission by telephone for the individual administration of the verbal portion of the creativity tasks.

To maintain complete records, it is necessary to have written verification of permission to test children on an individual basis. Please sign and return this permission form in the enclosed, stamped envelope. Again, your kind cooperation in facilitating this study is greatly appreciated.

Sincerely yours,

Anne Fishkin
Enrichment Teacher

My child, ____________________, was tested individually on the verbal portion of the Torrance Tests of Creative Thinking this fall with my consent. I understand that the posttest form of this instrument will be again administered individually in the late spring.

/_7 Permission was granted for individual testing on the verbal portion of the creativity tasks this fall. The posttest may also be given to my child in the spring.

/_7 Permission was granted for individual testing on the verbal portion of the creativity tasks this fall. I do <u>not</u> want my child to participate in the posttesting this coming spring.

____________________ ____________________
Date Parent's Signature

APPENDIX H

RATING SCALE OF WORK DONE BY THE TEAM:

CALCULATION OF EFFORT SCALE SCORE

RATING SCALE OF WORK DONE BY THE TEAM

Name____________________

Date__________ School__________

Instructions: We are interested in your opinion. Please do not tell others about your answers on this form. Please read the questions and think carefully before you answer each question.

1. How much work was completed by your team? Circle the answer which most clearly fits the amount of work you think your team did.

A	B	C	D	E
Very little work.	Some work.	Enough work.	Quite a lot of work.	A great amount of work.

2. List each member of your team. Using the lines below write their names in order from the person who spent the most time on the team's work to the person who spent the least time. (If possible, count time spent on assignments done at home as well as during team meetings.)

Name

Most time ____________________

____

____

____

____

Least time ____________________

3. List each member of your team. Using the lines below write their names in order from the most work to the least amount of work done by each person.)

Name

Most work ____________________

____

____

____

____

Least work ____________________

4. List each member of your team. Using the lines below write their names in order from the person who gave the most ideas to the person who gave the fewest ideas to the team.

Name

Most ideas ____________________

____

____

____

____

Least ideas ____________________

RATING SCALE OF WORK DONE BY THE TEAM:

CALCULATION OF EFFORT SCALE SCORE

The ranking was converted to a weighted ordinal measure by assigning a range of 1 to 3 for each question in order to control for variations in number of children on the teams. Teams varied from three on a team to six, with most teams being five. A child's rank for each question was assigned the value of that rank adjusted for the number of children on the team. For example, if there were five on the team, children, ranked from lowest to highest, were assigned scores of 1, 1.5, 2, 2.5, 3; if there were four on the team, the possible scores for any given question were 1, 1.7, 2.3, 3. These scores were summed for the three questions resulting in a possible range of 3 to 9 for the score of any individual's self or peer-ranking. The resultant scores for each team member were placed in rank order and the median peer/self-ranking score was calculated.

The obtained median score was adjusted by the value of the mean perceived amount of work performed by the team. The first question, "how much work was completed by your team?" would be answered for five different choices for increasing amount of work. These choices were assigned values from one to five. The perceived value of work completed, as the mean of each member's rating of the team's work, was added to the median of the child's peer/self-ranking score to yield the child's effort score. The possible range of scores, then, for effort was four to 14. Two examples of calculating the effort score are given below.

Step	Description	Child A	Child X
	Number on team	5	4
1	Self-ranking for question 4	BACDE	VXWY
2	Value of self-ranking for question 4	2.5	2.3
3	Median of self/peer-rankings for 3 questions	6.0	6.3
4	$\overline{X}$ team perception of work completed	3.6	4.0
5	Effort score (steps 4 + 5)	9.6	10.3

The range of possible effort scores was from 4 to 14. Effort scores ranged from 7.2 to 13.3. The median split was at 10.2, so that students whose effort scores were less than or equal to 10.2 were classed as OM-lo effort and students whose scores were greater than or equal to 10.3 were classed as OM-hi effort.

APPENDIX I

TEACHER RATING SCALE OF

TEAM EFFECTIVENESS

TEACHER RATING SCALE OF TEAM EFFECTIVENESS

Teacher's Initials__________
Date of Rating______________
Rating on which Future Problem? 1st 2nd
Number of previous years of experience with FPS? Circle one.
1st year 2nd 3rd 4th 5th
(No prior experience)

Student Name____________________________
Other students on this team:
____
____
____
____

	SD*	D	N	A	SA
1. This team had an overall good product.	1	2	3	4	5
2. Other members of the team listened to this student's ideas.	1	2	3	4	5
3. This team showed very poor effort in its work.	1	2	3	4	5
4. Team members were not cooperative.	1	2	3	4	5
5. This student seldom offered ideas to the team.	1	2	3	4	5
6. This student's work shows understanding of the fuzzy situation.	1	2	3	4	5
7. This student listened to other team members in a fair manner.	1	2	3	4	5
8. This student was critical of other members.	1	2	3	4	5
9. This student's contributions were a valuable asset to the team's final product.	1	2	3	4	5
10. This student's thoughts were not elaborated.	1	2	3	4	5
11. This student was not cooperative.	1	2	3	4	5
12. This student put effort into the team's work.	1	2	3	4	5
13. This student frequently tried to have decisions go in his direction regardless of the opinion of other team members.	1	2	3	4	5
14. This student showed mature skill when differences of opinion needed to be settled.	1	2	3	4	5
15. A disproportionate number of ideas came from members of the team other than this student.	1	2	3	4	5
16. In the Enrichment classes this child is productive.	1	2	3	4	5

*SD=strongly disagree; D =disagree; N=no opinion or not observed; A=agree; SA=strongly agree.

STUDENT EVALUATION OF OM Name____________________________

Date__________________ School__________________________

Directions: Check your answer to questions 1 and 2. For all other items, circle the number which most clearly fits the way you feel. Please read carefully and think before you make your choice.

1. I am on an OM team. (If you worked on a team which met only a few times and did not complete its work please also answer "yes.") Yes __ No __

2. I can be on an OM team next year. Yes __ Unsure __ No __

	Never	Sometimes	Often	Most of the time
3. I feel less creative than before I joined an OM team.	1	2	3	4
4. I feel that other people will use my ideas more often than before I joined an OM team.	1	2	3	4
5. I feel I can get more work done when I work with others than before I joined an OM team.	1	2	3	4
6. I feel less able to work together with others than before I joined an OM team.	1	2	3	4

For these items, circle the number which best fits the way you feel about how your team worked.

7. Most of the ideas came from 1 or 2 people while others did not give ideas.	1	2	3	4
8. Some people's ideas were ignored.	1	2	3	4
9. Team members did not criticize each other.	1	2	3	4
10. One or two persons tried to make other members use their ideas.	1	2	3	4
11. Some members did not help the team get its work done.	1	2	3	4
12. Work was shared by members of my team.	1	2	3	4
13. The team put very little effort into its work.	1	2	3	4
14. The team produced good solutions.	1	2	3	4

	Never	Sometimes	Often	Most of the time
15. Team members talked about different ideas and then picked the best ones.	1	2	3	4

For these items, circle the number which best fits the way you feel about how <u>you</u> worked on your team.

16. I offered my ideas to the team.	1	2	3	4
17. My ideas were ignored.	1	2	3	4
18. I ignored other people's ideas.	1	2	3	4
19. I did not criticize other members' ideas.	1	2	3	4
20. Other members of the team asked what I thought of their ideas.	1	2	3	4
21. I tried to make the team use my ideas.	1	2	3	4
22. I did not help the team get its work done.	1	2	3	4
23. I did not wait for others to tell me what to do.	1	2	3	4
24. I did my share of the team's work.	1	2	3	4
25. I did not help the team to complete its work.	1	2	3	4
26. I was an important member of the team in its work.	1	2	3	4
27. I am not glad I am on an OM team.	1	2	3	4
28. I would like to be on a team next year.	1	2	3	4

APPENDIX K

HOW MANY IDEAS? (IDEAS)

Date____________

HOW MANY IDEAS?

Name______________________________ Age______ Sex_____

School____________________________ Grade_________

Below is a list of ten statements which describe how people might see themselves. For each item circle the number which most clearly describes the way you feel about yourself. Please read carefully and think before you make your choice.

1.	In a group situation I am the one who provides a great many ideas.	Often	1	2	3	4	5	Seldom
2.	When I need to I find uncommon uses for everyday objects.	Often	1	2	3	4	5	Seldom
3.	When the first solution to a problem fails I am able to come up with other solutions.	Seldom	1	2	3	4	5	Often
4.	I come up with new ways to solve everyday problems.	Often	1	2	3	4	5	Seldom
5.	My friends consider me to be a creative person.	Seldom	1	2	3	4	5	Often
6.	My solutions or products for school projects are different from those of other students.	Seldom	1	2	3	4	5	Often
7.	Even when ideas are very different from each other, I can find relationships between them.	Often	1	2	3	4	5	Seldom
8.	When in a group discussion I suggest unusual ideas.	Seldom	1	2	3	4	5	Often
9.	I have more ideas than most of my friends.	Often	1	2	3	4	5	Seldom
10.	My thinking is very creative.	Seldom	1	2	3	4	5	Often

APPENDIX L

INTELLECTUAL ACHIEVEMENT RESPONSIBILITY

QUESTIONNAIRE (IAR)

Date______________

The IAR Questionnaire

Name____ Birthdate___ ______________________________

Grade_________ School_______________________ Sex_________________

GENERAL INSTRUCTIONS: This questionnaire describes a number of common experiences most of you have in your daily lives. These statements are presented one at a time, and following each are two possible answers. Read the description of the experience carefully, and then look at the two answers. Choose the one that most often describes what happens to you. Put a circle around the "A" or the "B" in front of that answer. Be sure to answer each question according to how you really feel.

If, at any time, you are uncertain about the meaning of a question, raise your hand and I shall come to your desk and try to explain it to you.

1. If a teacher passes you to the next grade, would it probably be

   A. because she liked you, or
   B. because of the work you did?

2. When you do well on a test at school, is it more likely to be

   A. because you studied for it, or
   B. because the test was especially easy?

3. When you have trouble understanding something in school, is it usually

   A. because the teacher didn't explain it clearly, or
   B. because you didn't listen carefully?

4. When you read a story and can't remember much of it, is it usually

   A. because the story wasn't well written, or
   B. because you weren't interested in the story?

5. Suppose your parents say you are doing well in school. Is this likely to happen

   A. because your school work is good, or
   B. because they are in a good mood?

6. Suppose you did better than usual in a subject at school. Would it probably happen

   A. because you tried harder, or
   B. because someone helped you?

7. When you lose at a game of cards or checkers, does it usually happen

   A. because the other player is good at the game, or
   B. because you don't play well?

8. Suppose a person doesn't think you are very bright or clever,

   A. can you make him change his mind if you try to, or
   B. are there some people who will think you're not very bright no matter what you do?

9. If you solve a puzzle quickly, is it

   A. because it wasn't a very hard puzzle, or
   B. because you worked on it carefully.

10. If a boy or girl tells you that you are dumb, is it more likely that they say that

   A. because they are mad at you, or
   B. because what you did really wasn't very bright?

11. Suppose you study to become a teacher, scientist, or doctor and you fail. Do you think this would happen

   A. because you didn't work hard enough, or
   B. because you needed some help, and other people didn't give it to you?

12. When you learn something quickly in school, is it usually

   A. because you paid close attention, or
   B. because the teacher explained it clearly?

13. If a teacher says to you, "Your work is fine," is it

   A. something teachers usually say to encourage pupils, or
   B. because you did a good job?

14. When you find it hard to work arithmetic or math problems at school, is it

   A. because you didn't study well enough before you tried them, or
   B. because the teacher gave problems that were too hard?

15. When you forget something you heard in class, is it

   A. because the teacher didn't explain it very well, or
   B. because you didn't try very hard to remember?

16. Suppose you weren't sure about the answer to a question your teacher asked you, but your answer turned out to be right. Is it likely to happen

A. because she wasn't as particular as usual, or
B. because you gave the best answer you could think of?

17. When you read a story and remember most of it, is it usually

A. because you were interested in the story, or
B. because the story was well written?

18. If your parents tell you you're acting silly and not thinking clearly, is it more likely to be

A. because of something you did, or
B. because they happen to feel cranky?

19. When you don't do well on a test at school, is it

A. because the test was especially hard, or
B. because you didn't study for it?

20. When you win at a game of cards or checkers, does it happen

A. because you play real well, or
B. because the other person doesn't play well?

21. If people think you're bright or clever, is it

A. because they happen to like you, or
B. because you usually act that way?

22. If a teacher didn't pass you to the next grade, would it probably be

A. because she "had it in for you," or
B. because your school work wasn't good enough?

23. Suppose you don't do as well as usual in a subject at school. Would this probably happen

A. because you weren't as careful as usual, or
B. because somebody bothered you and kept you from working?

24. If a boy or girl tells you that you are bright, is it usually

A. because you thought up a good idea, or
B. because they like you?

25. Suppose you became a famous teacher, scientist, or doctor. Do you think this would happen

A. because other people helped you when you needed it, or
B. because you worked very hard?

26. Suppose your parents say you aren't doing well in your school work. Is this likely to happen more

A. because your work isn't very good, or
B. because they are feeling cranky ?

27. Suppose you are showing a friend how to play a game and he has trouble with it. Would that happen

A. because he wan't able to understand how to play, or
B. because you couldn't explain it well?

28. When you find it easy to work arithmetic or math problems at school, is it usually

A. because the teacher gave you especially easy problems, or
B. because you studied your book well before you tried them?

29. When you remember something you heard in class, is it usually

A. because you tried hard to remember, or
B. because the teacher explained it well?

30. If you can't work a puzzle, is it more likely to happen

A. because you are not especially good at working puzzles, or
B. because the instructions weren't written clearly enough?

31. If your parents tell you that you are bright or clever, is it more likely

A. because they are feeling good, or
B. because of something you did?

32. Suppose you are explaining how to play a game to a friend and he learns quickly. Would that happen more often

A. because you explained it well, or
B. because he was able to understand it?

33. Suppose you're not sure about the answer to a question your teacher asks you and the answer you give turns out to be wrong. Is it likely to happen

A. because she was more particular than usual, or
B. because you answered too quickly?

34. If a teacher says to you, "Try to do better," would it be

A. because this is something she might say to get pupils to try harder, or
B. because your work wasn't as good as usual?

APPENDIX M

SEARS SELF-CONCEPT INVENTORY

Date______________

SEARS INVENTORY

Name________________________________ Age__________ Sex______

School________________________________ Grade__________

Some boys and girls have thought about the things they do and decided that the items on these pages were helpful in thinking about themselves. This is a chance for you to look at yourself and decide what your strong points are and what your weak points are. This is not a test; we expect everyone to have different answers--so be sure your answers show how you think about yourself. Your answers are private and will be kept in confidence.

Read each item and then answer the question: Compared with other boys and girls my age how do I rate now?

Find the line under whatever heading indicates your answer. (The words at the top show what the lines in each column stand for.) Mark an X on that line. Now go right ahead. Work as fast as you like.

	Excellent	Very good	Better than most	OK	Not so good
1. Being good at sports	___	___	___	___	___
2. Learning things rapidly	___	___	___	___	___
3. Making friends easily with my own sex	___	___	___	___	___
4. Having new, original ideas	___	___	___	___	___
5. Getting my school work done on time and not getting behind	___	___	___	___	___
6. Being able to read well	___	___	___	___	___
7. Being a good size and build for my age	___	___	___	___	___
8. Remembering what I've learned	___	___	___	___	___
9. Being willing for others to have their way sometimes	___	___	___	___	___
10. Solving problems in ways others haven't tried	___	___	___	___	___
11. Being confident; not shy or timid	___	___	___	___	___
12. Knowing how to do math	___	___	___	___	___
13. Being good at things that require physical skill	___	___	___	___	___
14. Being a good student	___	___	___	___	___
15. Being a leader--one to get things started with my own sex	___	___	___	___	___
16. Thinking up answers to problems--answers no one else has thought of	___	___	___	___	___
17. Being able to concentrate	___	___	___	___	___
18. Being interested in science; learning about things that scientists do	___	___	___	___	___
19. Being attractive, good looking	___	___	___	___	___

	Excellent	Very good	Better than most	OK	Not so good
20. Having brains for college	___	___	___	___	___
21. Making other people feel at ease	___	___	___	___	___
22. Learning about new things even when other people aren't interested—studying about things on my own	___	___	___	___	___
23. Getting a lot of fun out of life	___	___	___	___	___
24. Writing creative stories and poems	___	___	___	___	___
25. Being a good athlete	___	___	___	___	___
26. Being able to apply what I've learned	___	___	___	___	___
27. Having plenty of friends among my own sex	___	___	___	___	___
28. Seeing new ways of thinking about things and putting ideas together	___	___	___	___	___
29. Spending most of my time on my work, not goofing off	___	___	___	___	___
30. Having good handwriting even when I'm hurried	___	___	___	___	___
31. Being not too skinny, not too fat	___	___	___	___	___
32. Having brains	___	___	___	___	___
33. Being sensitive to what others are feeling	___	___	___	___	___
34. Being able to see things in my mind easily when I want to	___	___	___	___	___
35. Being able to change things when they don't suit me	___	___	___	___	___
36. Being able to spell correctly	___	___	___	___	___

	Excellent	Very good	Better than most	OK	Not so good
37. Enjoying games and sports	___	___	___	___	___
38. Being smart	___	___	___	___	___
39. Being active in social affairs with my own sex	___	___	___	___	___
40. Being interested in new things; excited about all there is to learn	___	___	___	___	___
41. Well organized; having materials ready when they're needed	___	___	___	___	___
42. Learning about people around the world and being interested in them	___	___	___	___	___
43. Having nice features (nose, eyes, etc.)	___	___	___	___	___
44. Knowing what to do to get the right answer to a problem	___	___	___	___	___
45. Being easy to get along with	___	___	___	___	___
46. Letting my imagination go when I want to	___	___	___	___	___
47. Enjoying myself in school	___	___	___	___	___
48. Doing well in art work, painting, or drawing	___	___	___	___	___

APPENDIX N

SAMPLE PROBLEMS FROM THE FUTURE PROBLEM SOLVING PROGRAM

Future Problem Solving
(Second of two problems for Grades 2 and 3)

Fuzzy Situation

As advances in medical technology have occurred, the people of America have begun to enjoy longer lifespans. Consequently there are now more older people in our society than ever before in history. People over 65 now make up about eleven percent (11%) of the American population. Some have referred to the influence of these older gray-haired Americans as the "graying" of America.

As older people retire, many find themselves living on incomes that are less than what they had when they were working. These incomes are also usually fixed, which means that they stay the same from year to year. Though inflation has been slowed down in recent years, it is still a problem for people on fixed incomes for it means that their limited amount of money is worth less and less each year.

In times past, older people often lived with their families, who cared for them in their later years. With the changes in American society and the fact that so many people no longer live in the same cities as their parents, caring for older relatives has become more difficult. It is becoming more frequent for older people, who can afford to do so, to move into Retirement Centers. These Centers have sprung up all over the country and are inhabited by people who no longer wish to live at home by themselves, or by those who are unable to do so. In addition to providing housing, meals and companionship, most retirement centers also offer entertainment and recreational opportunities. They also have health services nearby.

Despite the many positive aspects of the retirement homes, many of the people who reside there tend to become bored and listless. Many complain of feeling unwanted or unneeded, and many develop unexplained illnesses that doctors think are caused by their feeling of uselessness.

STEP I. Brainstorm all of the problems that you can think of. Be sure to number each problem.

STEP II. SELECTING THE PROBLEM.

From your ten best problems, select the one problem that you would like to try to solve. Write your problem so it begins with the words,"How might..?" or "In what ways might....?" Use only one verb and whatever conditions are necessary to describe the problem.

STEP III. BRAINSTORMING SOLUTION

Brainstorm solutions to only the problem you write in Step II. Be sure that each solutuion answers three of the six "journalistic" questions (who, what, when, where, why, and how). After you have completed Step III, go back and select your ten best solutions.

FUTURE PROBLEM SOLVING
(Second of two problems for Grades 4 and 5)

Name_______________________________

FUZZY SITUATION

PRACTICE PROBLEM TWO

FEEDING THE WORLD

"DROUGHT STRIKES . . . RECORD NUMBERS STARVING"

The above headline might have been found in newspapers during the 1970s, 1980s, or 1990s, but in this case, the year is 2000. Over 5,000 people in the world are dying every day from hunger or hunger-related diseases. Most of the starving people live in poor countries of the Third World, primarily the underdeveloped countries of Africa and Asia.

Progress in food production has stayed ahead of population growth, so that there is currently enough food to feed twice the world's population. But there is surplus food in some parts of the world, while other areas have shortages. These shortages, combined with an inability to distribute food fairly and efficiently, have contributed to world hunger.

Food shortages have increased because of a process known as desertification. Land which was used to produce limited amounts of food has been overused and will no longer grow food. Desert or "near-desert" conditions now occupy 50 percent of the Earth's surface.

Millions of people who lived and farmed in villages just outside of desert areas were forced to leave their homes as their land dried up during the 1990s. Many of these people chose to move to other villages, but the majority fled to cities where emergency relief was easier to get. Refugees in the cities are getting more to eat than refugees in the villages. Though the average diet is still below the recommended minimum requirements, a refugee in the city is consuming about 200 calories a day more than a refugee in a village. Medical help is also easier to get. Though the conditions are still far from ideal, a smaller percentage of refugees in the cities are dying of hunger.

Unfortuantely there has been a dramatic change in people's attitudes from the 1980s, when citizens in developed nations, encouraged by celebrities and rock stars, donated huge amounts of money to help feed the starving. By the end of the 20th century, these citizens had grown tired of donating food to countries, which were making few efforts to help their own poor and starving people. Journalists began to report on what was happening to emergency food supplies that were intended for the starving. Tons of food rotted on African docks. Some government officials seized emergency food shipments and resold them for personal profit. Emergency aid was not delivered because a signature was missing on official paperwork. Millions died as a result. Reports such as these discouraged donations, which are now 40 percent lower than they were during the peak period of the 1980s.

Nearly 2 million people, most of them children, will die this year from hunger. Use your problem solving skills as you tackle the issues related to this fuzzy situation.

FPSP

Part I. BRAINSTORMING POSSIBLE PROBLEMS OR DIFFICULTIES

Do not begin this section until after you have read the fuzzy situation and discussed it with the other members of your team. When you feel you understand the situation, brainstorm as many problems or difficulties as you can regarding the situation. List ONLY YOUR 20 BEST IDEAS, and number each.

FPSP

Part II. IDENTIFYING THE UNDERLYING PROBLEM

Select one of the problems and write it below. It should begin with the words "In what ways might..." or "How might..." Your problem should be written as clearly and specifically as possible. Write your problem below.

Part III. ALTERNATIVE SOLUTIONS

Brainstorm as many possible solutions as you can to the problem as you have defined it above. Record ONLY 20 OF YOUR TEAM'S SOLUTIONS below; please number your solutions.

FPSP

Part IV. CRITERIA FOR EVALUATING ALTERNATIVE SOLUTIONS

You have doubtless produced far more solutions than you need. Your problem now is to select your best alternative and make it better. List five criteria that you think are the most important for judging your solution. Make them specific, and choose ones that will differentiate your alternatives.

1.____

2.____

3.____

4.____

5.____

Part V. EVALUATING ALTERNATIVE SOLUTIONS

List below ten of your best alternative solutions and evaluate them according to your criteria. Rate each alternative on each criterion on a scale of 1 (poorest) to 10 (best).

ALTERNATIVE SOLUTIONS	CRITERIA					TOTAL
	1	2	3	4	5	
1.						
2.						
3.						
4.						
5.						
6.						
7.						
8.						
9.						
10.						

FPSP

Part VI. BEST SOLUTION

What is your best solution? Describe it in the space below, answering some of the following questions: Who will carry out the solution? Where will it be carried out? When will it be carried out? How will the action to be taken accomplish the solution? What obstacles will need to be overcome?

APPENDIX O

INSTRUCTIONS TO PRECEDE WARM-UP ACTIVITIES FOR POSTTESTING AND SAMPLE WARM-UP ACTIVITIES FOR CREATIVE AND AFFECTIVE TESTING SESSIONS

INSTRUCTIONS TO PRECEDE WARM-UP ACTIVITIES

FOR POSTTESTING

This year in Enrichment classes we are working on a special study to look at how much you grow during the year. At the beginning of the year we asked for your parents' permission for you to work on the special activities for the study. During the next few weeks, part of the class time will be used for those of you with permission to do these activities. Those without permission to be in the study, or who started class after we began this project, will be able to work on a choice of quiet centers or on a critical thinking workbook during the special study times. We will all do the warm-up activities together.

Warm-Up (Figural)

(Pass out Renzulli, 12B for posttest (Figural Activity Sheet). Read the instructions on activity sheet. Add the following:) In the next few minutes try to come up with at least three answers for the warm-up activity sheet. Try to think of as many ideas, and as many different kinds of ideas that it reminds you of. Feel free to add to, or elaborate to your drawing. We will then briefly share some of our favorite ideas. (Allow two minutes for drawing; two minutes for sharing of ideas.)

Administration of TTCT Figural

(Pass out Figural Booklets. Read first paragraph on page 6: "I believe . . .") Please fill out the information asked for carefully, including today's date, your first and last name, and you can abbreviate your school if you wish. (Read second paragraph on page 6.) "One of the things . . . ability to think of ideas. As you remember, we took a measurement in the fall and today we are measuring your ability to think of ideas in these ways again. We want to get as accurate a measurement today as we can. So use your besting thinking cap and do your best. (Continue with administration instructions as in the manual.)

(After they have completed Circles, Test 3, ask children to check activities #2 and #3 to make sure they have a title for each of their drawings. Pass booklets up and take a short break.)

Similes

(Give a short break. Administer on same day as TTCT Figural if at all possible. No warm-up.) We will now do another creative activity, only this time we will write down our answers.

(Kathleen or Barbara: If at all possible while children are working on Similes, please check to make sure they have titles for all of their drawings, and that the titles communicate well enough so that I can score them and give them their full credit. Thanks.)

Warm-Up (Verbal)

(Pass out Renzulli B posttest verbal activity sheet.) Last week in class we did some creative thinking activities as part of this year's special study. Today we will again do some using a different kind of thinking.

(Read the instructions on the activity sheet. Add the following:) In the next few minutes try to come up with at least four answers for the warm-up activity sheet. Try to think of as many ideas, and as many different kinds of ideas that you can. Feel free to describe things in different ways. We will then briefly share some of our favorite ideas. (Allow two minutes for writing; two minutes for sharing of ideas.)

Administration of TTCT Verbal

The activities in the booklet will call for all the imagination and thinking ability you have. So I hope you will put on your best thinking cap and that you will enjoy yourself. (Pass out the booklets. Ask them to carefully complete all of the requested information. They may abbreviate their school if they like.) We want to get as accurate a measurement today as we can. So give it your best thought so as to do your best. (Begin reading instructions on page 5.) "The activities in this booklet . . ."

Note: Activity 4: (sketch of a stuffed toy monkey . . . $9 or 10.) Activity 6 "Unusual Questions" is again omitted.

(Individual testing: Re-administer the same warm-up activity. After sharing, brainstorm with the child to help remember some ideas shared in class. The examiner may offer some ideas to help the child's discussion with her. Sample ideas are included on the attached page for Verbal B warm-up.)

Affective Activities Warm-Up (To precede pink form, "How Many Ideas?")

Within the past few weeks we did some activities to measure your ability to think of ideas. Today we will do two activities that measure the way we feel about ourselves.

Affective Warm-Up Activity

(Distribute warm-up activity sheet "How I Do Things." Give class three-five minutes to complete the activity and approximately another three minutes to share their responses if they choose to do so.)

WARM-UP ACTIVITIES FOR AFFECTIVE PRETESTING

Name______________________________

ME, MYSELF, AND THEM

WHAT I KNOW ABOUT MYSELF HELPS ME TO UNDERSTAND OTHERS, WHAT I DISCOVER ABOUT OTHERS HELPS ME TO UNDERSTAND MYSELF. CHECK THE ANSWER TO EACH QUESTION THAT BEST FITS YOU.

1. When I look at myself I see that I:

___Like myself as I am
___Wish that I were different
___Wish I were like a friend

2. When I play with others I generally:

___Expect others to do as I say
___Expect to be told what to do
___Sometimes lead, sometimes follow

3. With my friends I am generally:

___Lots of fun
___Considerate
___Understanding
___A good sport
___A sore loser
___Not always considerate

4. When it comes to laughter, I find I:

___Tend to be serious
___Have a sense of humor
___Am sometimes silly
___Laugh easily
___Need a really good laugh
___Don't like to be laughed at

5. I am generally (check no more than 3):

___Cheerful
___Serious
___Moody
___Thoughtful
___Responsible
___Patient
___Impatient
___Enthusiastic
___Short-tempered

Finish this paragraph: I am the kind of person who______________________________

____

____

____

from Jeanne Heiberg, 1977

APPENDIX P

SUPPLEMENTARY RESULTS REPRESENTED IN TABLES AND FIGURES

Table P-1

Within Cells Correlations of the Ten Creative Dependent Variables Showing Relationships of Pretest and Posttest Measures[a]

	Similes	Verbal TTCT			Figural TTCT					
		Fluency	Flexi-bility	Origi-nality	Fluency	Origi-nality	Titles	Elabo-ration	Closure	Creative Strengths
Similes	--									
Verbal TTCT										
Fluency	-.13[b] -.06	--								
Flexibility	-.14 .03	.91** .79**	--							
Originality	-.19* -.10	.95** .84**	.92** .76**	--						
Figural TTCT										
Fluency	.08 -.10	.27** .12	.17 .14	.16 .12	--					
Originality	.15 -.08	.16 .07	.06 .12	.11 .06	.60** .46**	--				
Titles	.16 -.14	.02 -.09	-.01 -.12	.00 -.02	-.00 -.03	-.03 -.08	--			
Elaboration	.25** .03	.13 .00	.20* .04	.14 .05	.13 .08	.11 -.13	.11 .21*	--		
Closure	.14 -.22*	.02 .16	-.09 .06	.00 .20*	.19* -.04	.26** .09	.01 .18	.09 .25**	--	
Creative Strengths	.36** -.04	-.07 .17	-.08 .09	-.10 .11	.00 .19*	.21* .13	.34** .22*	.53** .42**	.30** .10	--

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation. $n = 116$.

[a]For this analysis, scores for the 17 students whose standard scores were greater than 160 are reported as obtained.
[b]Pretest intercorrelations are reported above posttest intercorrelations.

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

Table P-2

Cell Means and Standard Deviations of Ten Creative Dependent Variables and Odyssey of the Mind (OM) Effort[a]

		Similes		Verbal TTCT Dependent Variables					
				Fluency		Flexibility		Originality	
OM Effort	n	X̄	SD	X̄	SD	X̄	SD	X̄	SD
Pretest Means									
Non-OM	51	37.92	13.64	131.41	30.71	141.78	29.18	133.78	31.15
OM-lo	32	46.53	11.27	126.25	25.85	133.53	16.99	126.72	23.48
OM-hi	33	50.24	21.21	127.82	27.61	133.91	22.48	128.00	24.13
Posttest Means									
Non-OM	51	47.67	16.16	119.82	17.85	132.86	19.18	108.86	14.74
OM-lo	32	49.44	14.41	116.50	20.41	131.13	20.64	107.09	19.93
OM-hi	33	61.12	18.21	126.64	21.57	140.30	24.20	115.76	17.36

		Figural TTCT Dependent Variables											
		Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
		X̄	SD	X̄	SD	X̄	SD	X̄	SD	X̄	SD	X̄	SD
Pretest Means													
Non-OM	51	98.39	16.22	114.84	20.32	110.77	17.39	124.78	19.81	105.88	13.01	8.49	2.94
OM-lo	32	104.31	14.30	124.63	17.83	109.00	13.67	133.69	20.36	111.97	9.33	9.47	2.21
OM-hi	33	105.64	18.94	124.64	20.24	111.91	21.05	134.58	20.85	110.46	11.16	10.15	3.25
Posttest Means													
Non-OM	51	107.69	15.21	137.77	25.40	112.35	18.03	116.29	19.85	105.82	17.01	8.84	2.82
OM-lo	32	110.03	14.91	141.41	27.11	113.63	17.95	120.88	22.39	113.59	13.82	9.75	2.98
OM-hi	33	107.76	18.58	146.91	24.83	116.15	20.66	121.64	19.69	117.70	17.26	11.24	3.35

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation.

[a]For this analysis, scores for the 17 students whose standard scores were greater than 160 are reported as obtained.

Table P-3

Summary of Mixed Model Creative MANOVA for Odyssey of the Mind (OM) Effort: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Ten Creative Dependent Variables[a]

Source	Multivariate Global F	Stepdown Fs/Univariate Fs — Dependent Variables: Similes	Verbal TTCT Fluency	Verbal TTCT Flexi-bility	Verbal TTCT Origi-nality
Between Subjects Analyses					
OM	1.68*[b]	3.44*[c] 8.84***	1.52 .70	.02 .70	.18 .69
Repeated Subjects Analyses					
Time	29.76***	25.22*** 28.92***	7.38** 11.36***	1.27 1.77	41.38*** 74.01***
OM x Time	1.53	2.53 2.34	.97 1.70	5.33** 4.69**	.14 2.61

Source	Figural TTCT Fluency	Figural TTCT Origi-nality	Figural TTCT Titles	Figural TTCT Elabo-ration	Figural TTCT Closure	Creative Strengths
Between Subjects Analyses						
OM	1.19 1.19	1.65 2.67	.39 .36	1.78 2.56	4.80** 7.54***	2.10 7.15***
Repeated Subjects Analyses						
Time	13.44*** 13.44***	63.46*** 86.72***	3.21 2.60	11.11*** 27.31***	.73 2.02	2.30 3.03
OM x Time	1.54 1.54	1.03 .68	.21 .26	.63 .54	1.41 1.55	1.36 .62

[a]For this analysis scores for the 17 students whose standard scores were greater than 160 are reported as obtained.
[b]Pillai's criterion for statistical inference. [c]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001

Table P-4

Cell Means and Standard Deviations of Ten Creative Dependent Variables,

Odyssey of the Mind (OM) Effort and Experience in Enrichment Program[a]

Experience in Enrichment Program	OM Effort	n	Similes		Verbal TTCT Dependent Variables: Fluency		Flexibility		Originality	
			$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Pretest Means										
< 13 Months	Non-OM	18	31.22	9.01	131.33	22.84	140.83	17.73	132.83	25.18
	OM-lo	11	48.18	11.17	126.73	18.66	132.27	13.06	127.82	18.40
	OM-hi	11	41.27	22.49	131.82	21.61	137.36	15.34	129.91	20.91
≥ 13 Months	Non-OM	33	41.58	14.44	125.67	22.85	133.73	19.70	127.09	19.90
	OM-lo	21	45.67	11.50	122.86	20.22	133.33	17.24	123.76	19.60
	OM-hi	22	54.73	19.53	120.77	18.93	127.68	15.48	122.96	16.96
Posttest Means										
< 13 Months	Non-OM	18	43.33	17.10	122.83	19.73	140.40	16.34	113.50	16.36
	OM-lo	11	51.09	13.03	115.46	18.17	132.09	15.38	105.82	17.64
	OM-hi	11	54.27	20.98	130.00	22.74	139.55	22.39	121.36	21.05
≥ 13 Months	Non-OM	33	50.03	15.38	118.18	16.82	128.15	18.49	106.33	13.36
	OM-lo	21	48.57	15.32	116.48	20.51	128.91	19.96	107.05	19.20
	OM-hi	22	64.55	16.08	124.27	19.96	136.41	19.60	112.96	14.95

			Figural TTCT Dependent Variables: Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
			$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Pretest Means														
< 13 Months	Non-OM	18	96.17	14.90	109.61	21.63	114.94	18.10	116.33	19.21	102.06	14.04	7.67	2.54
	OM-lo	11	102.46	13.23	124.55	21.33	109.18	16.95	126.91	22.56	108.55	9.89	8.55	1.97
	OM-hi	11	103.00	20.13	119.36	18.95	108.00	25.04	131.73	12.63	110.09	9.28	9.64	3.23
≥ 13 Months	Non-OM	33	99.61	17.00	117.52	18.85	108.49	16.83	129.30	18.69	107.97	12.13	8.94	3.07
	OM-lo	21	105.29	15.05	124.19	15.00	108.91	12.07	136.14	16.89	113.76	8.73	9.95	2.22
	OM-hi	22	106.96	18.65	127.00	20.27	113.86	19.09	134.23	21.72	110.64	12.20	10.41	3.31
Posttest Means														
< 13 Months	Non-OM	18	107.33	11.48	125.50	23.15	113.67	20.01	111.94	18.93	100.94	15.91	7.11	2.25
	OM-lo	11	108.46	11.87	141.09	22.14	116.18	21.62	112.27	21.91	110.27	12.78	8.82	3.09
	OM-hi	11	102.18	16.39	131.36	20.02	115.55	20.32	120.64	13.01	112.73	15.82	10.46	2.54
≥ 13 Months	Non-OM	33	107.88	17.06	140.39	17.80	111.46	16.56	118.67	20.23	108.49	17.23	9.79	2.67
	OM-lo	21	110.86	16.49	135.91	22.62	112.29	16.13	125.38	21.80	115.33	14.33	10.24	2.88
	OM-hi	22	110.55	19.33	147.55	13.78	116.18	20.65	122.14	22.57	120.18	17.77	11.64	3.67

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation.

[a] All standard scores set at ≤ 160.

Table P-5

Summary of Mixed Model Creative MANOVA for Odyssey of the Mind (OM) Effort and Experience in the Enrichment Program: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Ten Creative Dependent Variables[a]

Source	Multivariate Global F	Stepdown Fs/Univariate Fs — Dependent Variables: Similes	Verbal TTCT: Fluency	Verbal TTCT: Flexibility	Verbal TTCT: Originality
Between Subjects Analyses					
OM	2.01**[b]	4.08*[c] 9.27***	1.58 .75	.57 .37	1.05 .75
Experience in Enrichment Program	3.13**	1.72 6.50**	2.54 2.14	4.49* 4.35*	.00 3.09
Experience x OM	1.21	2.66 2.40	.18 .29	3.77* .67	.16 .34
Repeated Subjects Analyses					
Time	31.43***	26.03*** 28.37***	7.54** 7.62**	.73 .14	44.72*** 86.55***
OM x Time	1.24	2.56 2.29	1.36 2.66	1.93 3.74*	.35 3.27*
Experience x Time	.73	.25 .60	.19 .75	2.64 .18	.18 .02
Experience x OM x Time	.71	.01 .12	.19 .13	1.95 1.07	1.22 .32

Source	Stepdown Fs/Univariate Fs — Dependent Variables — Figural TTCT: Fluency	Originality	Titles	Elaboration	Closure	Creative Strengths
Between Subjects Analyses						
OM	1.16 1.16	1.34 2.84	.45 .36	1.73 2.33	5.10** 7.69***	2.65 7.52***
Experience in Enrichment Program	1.63 1.63	4.22 5.87*	.13 .29	5.76* 6.45*	4.45* 7.88**	4.11* 12.06***
Experience in OM	.23 .23	3.00 2.01	.60 .60	.99 .70	.28 .16	.24 .43
Repeated Subjects Analyses						
Time	13.17*** 13.17***	63.94*** 84.28***	2.66 2.41	11.38*** 25.86***	.55 1.98	3.99* 3.00
OM x Time	1.51 1.51	1.13 1.14	.20 .26	.53 .36	1.41 1.51	1.39 .59
Experience x Time	.01 .01	1.46 1.03	.00 .02	.10 .23	.40 .54	2.16 1.30
Experience x OM x Time	.35 .35	1.23 .95	.58 .54	.45 .48	.32 .28	.87 .43

[a]All standard scores set at ≤ 160. [b]Pillai's criterion for statistical inference.
[c]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001.

Table P-6

Cell Means and Standard Deviations of Creative Dependent Variables, Odyssey of the Mind (OM) Effort, and Prior Experience on an OM Team[a]

Prior Experience on an OM Team	OM Effort	n	Similes X̄	Similes SD	Verbal TTCT Fluency X̄	Fluency SD	Flexibility X̄	Flexibility SD	Originality X̄	Originality SD
Pretest Means										
No Prior Experience	Non-OM	39	34.33	12.38	131.51	22.21	139.26	18.41	132.44	22.02
	OM-lo	12	45.58	12.04	125.75	19.01	132.58	14.16	128.33	17.80
	OM-hi	10	38.30	20.45	128.70	20.75	134.60	14.50	126.90	19.37
Experience on an OM Team	Non-OM	12	49.58	11.02	115.17	20.78	126.42	18.98	118.33	18.04
	OM-lo	20	47.10	11.06	123.25	20.19	133.20	16.94	123.25	19.88
	OM-hi	23	55.44	19.75	122.61	20.17	129.30	16.49	124.57	18.28
Posttest Means										
No Prior Experience	Non-OM	39	45.72	16.80	121.92	17.40	135.92	16.33	110.74	15.69
	OM-lo	12	48.83	14.22	117.83	16.60	133.08	15.63	105.92	16.55
	OM-hi	10	54.50	19.28	129.70	22.92	137.70	22.47	119.60	21.31
Experience on an OM Team	Non-OM	12	54.00	12.47	113.00	18.30	121.25	21.54	102.75	9.11
	OM-lo	20	49.80	14.87	115.10	21.32	128.15	19.91	107.05	19.83
			64.00	17.36	124.65	20.08	137.35	19.78	114.09	15.59

Figural TTCT Dependent Variables

Prior Experience on an OM Team	OM Effort	n	Fluency X̄	Fluency SD	Originality X̄	Originality SD	Titles X̄	Titles SD	Elaboration X̄	Elaboration SD	Closure X̄	Closure SD	Creative Strengths X̄	Creative Strengths SD
Pretest Means														
No Prior Experience	Non-OM	39	97.54	17.05	113.31	21.83	111.97	18.06	122.74	19.74	105.33	13.69	8.00	2.72
	OM-lo	12	105.33	12.95	126.25	19.87	108.43	14.98	130.50	23.75	111.50	10.28	9.00	2.30
	OM-hi	10	104.30	21.23	119.60	19.96	112.60	27.06	133.20	15.83	111.40	10.46	9.70	3.34
Experience on an OM Team	Non-OM	12	101.17	13.42	119.33	12.12	106.83	15.01	131.17	18.94	107.67	10.82	10.08	3.15
	OM-lo	20	103.70	15.35	123.15	15.62	109.35	13.21	134.45	16.37	112.25	8.99	9.75	2.17
	OM-hi	23	106.00	18.35	126.57	19.92	111.61	18.56	132.48	20.55	110.04	11.66	10.35	3.27
Posttest Means														
No Prior Experience	Non-OM	39	109.72	13.28	133.18	21.56	112.28	18.96	116.62	20.18	102.13	14.42	8.36	2.78
	OM-lo	12	112.25	14.80	141.75	21.24	109.33	20.49	112.83	21.31	111.42	12.81	8.58	2.47
	OM-hi	10	98.60	14.16	132.60	18.67	112.70	22.02	120.50	14.90	114.50	16.18	9.80	3.16
Experience on an OM Team	Non-OM	12	101.08	19.49	141.50	17.93	112.08	13.44	115.25	19.57	117.83	19.77	10.42	2.43
	OM-lo	20	108.70	15.19	135.25	23.00	116.20	16.25	125.70	22.13	114.90	14.56	10.45	3.10
	OM-hi	23	111.74	19.11	156.30	15.77	117.39	19.74	122.13	21.73	119.09	17.88	11.87	3.29

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at ≤ 160.

Table P-7

Summary of Mixed Model Creative MANOVA for Odyssey of the Mind (OM) Effort and Prior Experience on an OM Team: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Ten Creative Dependent Variables[a]

Source	Multivariate Global F	Similes	Verbal TTCT Fluency	Verbal TTCT Flexibility	Verbal TTCT Originality
		Stepdown Fs/Univariate Fs — Dependent Variables			
Between Subjects Analyses					
OM	1.30[b]	2.43[c] 4.20*	1.48 1.04	.36 .43	.84 1.00
Prior Experience on an OM Team	4.76***	9.91** 22.06***	2.13 3.74	.26 4.48*	3.13 3.08
Prior Experience x OM	.98	2.02 1.91	.32 .74	1.43 1.48	.54 .79
Repeated Subjects Analyses					
Time	31.16***	25.86*** 28.97***	7.14** 7.64**	.80 .14	45.88*** 87.53***
OM x Time	1.01	2.47 2.13	1.16 2.12	2.24 3.56*	.21 2.37
Prior Experience x Time	1.50	.96 2.75	1.00 1.73	.00 .33	.54 2.63
Prior Experience x OM x Time	1.16	.16 .43	.24 .35	2.49 .63	2.10 .58

Source	Figural TTCT Fluency	Figural TTCT Originality	Figural TTCT Titles	Figural TTCT Elaboration	Figural TTCT Closure	Creative Strengths
	Stepdown Fs/Univariate Fs — Dependent Variables					
Between Subjects Analyses						
OM	.92 .92	.67 1.32	.33 .28	.66 .99	2.29 3.50*	1.35 2.76
Prior Experience on an OM Team	.56 .56	4.52* 4.58*	.34 .19	3.16 4.19*	8.93** 13.47***	9.04** 21.45***
Prior Experience x OM	1.26 1.26	1.70 1.69	.31 .44	.53 .38	1.39 1.35	.35 .29
Repeated Subjects Analyses						
Time	13.96*** 13.96***	61.34*** 82.84***	2.69 2.42	11.94*** 26.31***	.83 2.07	3.42 3.00
OM x Time	1.04 1.04	1.27 1.19	.01 .02	.38 .26	.54 .64	.74 .40
Prior Experience x Time	1.37 1.37	.30 .01	2.00 2.04	.40 .23	5.54* 6.07*	2.73 1.62
Prior Experience x OM x Time	3.55* 3.55*	.95 .43	.00 .00	1.09 1.56	1.00 .72	.03 .42

[a]All standard scores set at ≤ 160. [b]Pillai's criterion for statistical inference.
[c]Stepdown Fs are above univariate Fs.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table P-8

Summary of the Creative Stepwise Multiple Regression Analyses: Contributions of 26 Predictor Variables for Ten Regression Analyses for the Ten Creative Dependent Variables[a]

	Posttest Dependent Variables											
	Similes			Verbal TTCT								
Predictor Variables				Fluency			Flexibility			Originality		
Statistics for Prediction Equation for all Predictor Variables												
Multiple R		.56			.95			.92			.94	
R^2		.31***[b]			.91***			.85***			.88***	
Adjusted R^2		.30			.90			.84			.88	
F		25.36***			132.60***			120.77***			136.12***	
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Pretest Variables												
Similes	.48	.46	7.36**									
Verbal Fluency										-.11	-.14	8.62**
Verbal Flexibility												
Verbal Originality				.16	.17	16.11***						
Figural Fluency												
Figural Originality												
Figural Titles												
Figural Elaboration				-.10	-.10	9.03**				.10	.11	10.64**
Figural Closure												
Creative Strengths	.54	.51	41.07***									

(table continues)

Table P-8 (continued)

Predictor Variables	Posttest Dependent Variables											
	Similes			Verbal TTCT								
				Fluency			Flexibility			Originality		
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Posttest Variables												
Similes												
Verbal Fluency							.80	.82	379.05***	.64	.74	310.09***
Verbal Flexibility				.96	.94	414.75***						
Verbal Originality				.50	.44	43.93***						
Figural Fluency												
Figural Originality												
Figural Titles												
Figural Elaboration												
Figural Closure												
Creative Strengths							.64	.11	6.55**			
Control Variables												
Months in Program												
Years in OM												
Grade				16.41	.41	3.31*	24.33	.62	13.18***	8.17	.08	19.66***
Teacher				.69	.04	2.46^{c}	9.60	.24	1.86			
OM				-19.86	-.48	2.90				-20.55	-.58	5.55**
Sex												

(table continues)

Table P-8 (continued)

Predictor Variables	Posttest Dependent Variables: Figural TTCT: Fluency			Originality			Titles			Elaboration			Closure[d]			Creative Strengths		
Statistics for Prediction Equation for all Predictor Variables																		
Multiple R		.55			.60			.56			.69			.74			.72	
R^2		.30***			.36***			.31***			.48***			.55***			.52***	
Adjusted R^2		.28			.34			.28			.45			.53			.50	
F		12.09***			20.84***			9.97***			16.64***			26.90***			24.31***	
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Pretest Variables																		
Similes																		
Verbal Fluency							.17	.20	5.63*									
Verbal Flexibility																		
Verbal Originality																		
Figural Fluency	.17	.18	3.98*										.24	.23	11.71***			
Figural Originality				.48	.46	31.16***												
Figural Titles							.27	.25	8.62**									
Figural Elaboration										.30	.29	12.80***						
Figural Closure							-.37	-.24	8.40**									
Creative Strengths																.26	.24	11.07***

(table continues)

Table P-8 (continued)

Predictor Variables	Posttest Dependent Variables: Figural TTCT — Fluency			Originality			Titles			Elaboration			Closure[d]			Creative Strengths		
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Posttest Variables																		
Similes																		
Verbal Fluency	.23	.27	10.90***															
Verbal Flexibility							-.20	-.20	4.67*									
Verbal Originality																		
Figural Fluency				.39	.31	14.53***				.20	.16	4.00*						
Figural Originality	.31	.39	20.92***							-.17	-.17	-5.10*						
Figural Titles																.04	.22	9.51**
Figural Elaboration																.08	.51	40.15***
Figural Closure																		
Creative Strengths							2.09	.36	16.71***									
Control Variables																		
Months in Program																		
Years in OM													6.92	.41	22.40***			
Grade										-6.07	-.14	7.97**	-33.72	-1.00	60.05***	-2.25	-.35	22.81***
Teacher																		
OM													-6.94	-.20	6.79**			
Sex	7.66	.23	7.63**	-10.15	-.24	9.45**				10.26	.24	10.81***				-1.37	-.21	9.62**

Note. Three of the control variables, Grade, Teacher, and OM, were coded into dummy variables to satisfy the assumptions required for regression analyses (Cohen & Cohen, 1983; Tabachnick & Fidell, 1983).
Note. The predictor variables in the table with negative signs are suppressor variables, in actuality carrying very low correlation with the dependent variables in the equation. The suppressor variables (Grade, OM, and Sex) suppress variance in other possible predictor variables which are irrelevant to the prediction of the dependent variable (Tabachnick & Fidell, 1983).

[a]All standard scores set at ≤ 160.
[b]The reported F ratios resulting from stepwise regression represent an inflated alpha Type I error. Greater confidence can be placed in the significance levels associated with Multiple R^2 (Tabachnick & Fidell, 1983).
[c]The dummy coded variables may have been included in the equation because one level of that variable remained in the equation and the other portion was removed, thus yielding a non-significant total Fchange value.
[d]Standard Forward Regression for Closure because of large values for suppressor variables.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table P-9

Cell Means and Standard Deviations of Ten Creative Dependent Variables, Odyssey of the Mind (OM) Effort, and Grade[a]

			Similes		Verbal TTCT Dependent Variables					
					Fluency		Flexibility		Originality	
OM Effort	Grade	n	X̄	SD	X̄	SD	X̄	SD	X̄	SD
Pretest Means										
Non-OM	2+3	27	32.85	10.58	140.11	18.87	147.56	14.47	139.85	20.94
	4	15	40.40	16.57	114.33	19.83	123.47	14.78	118.87	17.28
	5	9	49.00	9.03	112.56	16.89	123.56	17.70	114.00	13.76
OM-lo	2+3	8	46.63	11.86	143.13	10.12	149.00	7.27	143.88	10.79
	4	16	43.44	11.54	116.31	17.21	127.25	14.82	120.06	17.10
	5	8	52.63	8.50	121.00	19.52	128.38	13.35	116.63	17.68
OM-hi	2+3	7	30.43	19.33	145.71	19.29	146.14	16.42	142.86	22.73
	4	14	55.43	16.14	120.57	17.28	131.07	13.77	125.71	15.36
	5	12	55.75	21.91	116.58	16.11	121.83	11.19	114.50	9.68
Posttest Means										
Non-OM	2+3	27	44.33	16.76	122.56	17.89	140.96	15.21	116.00	14.98
	4	15	49.13	16.61	120.80	15.24	126.80	12.55	99.60	8.43
	5	9	55.22	11.48	110.00	20.21	116.44	23.06	102.89	11.42
OM-lo	2+3	8	43.00	17.79	128.38	24.27	145.50	18.50	126.88	22.86
	4	16	51.25	15.07	111.13	14.08	124.50	12.60	96.56	6.40
	5	8	52.25	7.21	113.88	20.72	125.50	20.90	106.50	13.73
OM-hi	2+3	7	51.29	11.43	137.57	20.87	151.14	15.99	133.14	21.40
	4	14	66.57	13.63	126.79	23.67	136.86	19.90	110.29	14.32
	5	12	60.50	23.95	118.83	14.51	130.17	20.27	112.00	11.78

			Figural TTCT Dependent Variables											
			Fluency		Originality		Titles		Elaboration		Closure		Creative Strengths	
			X̄	SD	X̄	SD	X̄	SD	X̄	SD	X̄	SD	X̄	SD
Pretest Means														
Non-OM	2+3	27	95.52	14.43	112.22	22.34	111.04	20.34	122.30	20.64	104.96	14.47	8.04	2.98
	4	15	102.80	21.39	118.87	20.85	111.73	13.63	131.53	19.98	106.73	13.13	8.80	2.65
	5	9	99.67	10.05	115.33	8.41	108.33	14.58	120.67	14.73	107.22	8.33	9.33	3.32
OM-lo	2+3	8	100.13	8.83	126.50	8.40	99.25	13.90	138.25	27.18	113.25	13.58	9.88	2.64
	4	16	106.19	14.97	123.38	20.36	112.13	12.71	133.69	16.61	113.25	7.98	9.19	2.26
	5	8	104.75	17.87	124.00	18.07	112.50	11.95	126.25	14.68	108.13	6.51	9.63	1.85
OM-hi	2+3	7	93.29	12.62	120.57	22.44	96.71	12.75	132.57	16.14	116.86	9.42	9.43	2.57
	4	14	108.00	20.62	125.50	19.83	118.64	18.63	141.43	19.89	106.21	10.31	9.29	3.34
	5	12	110.08	18.11	125.50	19.96	112.92	24.16	124.50	16.47	111.67	11.74	11.58	3.23
Posttest Means														
Non-OM	2+3	27	108.82	12.36	132.85	22.91	110.93	18.00	116.74	23.00	103.26	15.29	7.85	2.43
	4	15	107.60	16.27	137.13	18.77	109.27	14.84	115.47	11.28	98.93	10.31	9.47	2.85
	5	9	104.44	21.66	138.67	19.16	121.11	20.77	116.33	22.81	125.00	18.37	10.78	2.82
OM-lo	2+3	8	111.38	13.45	138.88	23.78	110.00	19.49	116.00	27.17	107.25	15.83	7.75	3.77
	4	16	112.19	15.36	142.19	16.55	113.19	17.86	118.81	21.48	109.50	7.71	10.25	2.70
	5	8	104.38	15.78	127.50	29.51	118.13	18.04	129.88	19.08	128.13	11.93	10.75	1.83
OM-hi	2+3	7	95.29	11.46	132.29	24.57	108.86	14.82	126.00	7.05	106.86	7.11	10.29	2.43
	4	14	115.71	21.90	149.79	11.42	120.14	21.65	115.93	20.22	105.43	6.95	11.14	2.96
	5	12	105.75	13.47	139.00	16.47	115.25	21.48	125.75	23.36	138.33	7.18	11.92	4.23

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at $\leq$ 160.

Table P-10

Summary of Mixed Model Creative MANOVA for Odyssey of the Mind (OM) Effort and Grade: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Ten Creative Dependent Variables[a]

Source	Multivariate Global F	Stepdown Fs/Univariate Fs: Dependent Variables: Similes	Verbal TTCT: Fluency	Verbal TTCT: Flexi-bility	Verbal TTCT: Origi-nality
Between Subjects Analyses					
OM	1.81*[b]	2.61[c] 5.16**	1.63 2.34	.61 2.40	3.68* 4.25*
Grade	10.82***	6.87** 12.74***	23.77*** 15.71***	9.78*** 29.50***	11.10*** 33.19***
Grade x OM	1.25	1.67 1.45	1.54 .50	1.52 .43	2.06 .64
Repeated Subjects Analyses					
Time	42.04***	28.36*** 29.49***	5.86* 8.76**	.16 .13	72.88*** 90.64***
OM x Time	1.49	2.80 2.49	1.27 2.12	3.63* 3.08*	2.67 2.24
Grade x Time	9.67***	.58 1.46	6.93** 10.19***	4.08* 1.28	24.55*** 4.83**
OM x Grade x Time	.97	1.17 1.28	1.09 .56	.84 .53	.15 .40

Source	Stepdown Fs/Univariate Fs: Dependent Variables: Figural TTCT: Fluency	Figural TTCT: Origi-nality	Figural TTCT: Titles	Figural TTCT: Elabo-ration	Figural TTCT: Closure	Creative Strengths
Between Subjects Analyses						
OM	.44 .44	1.55 1.82	.25 .25	1.60 1.94	4.02* 6.45**	1.56 3.99*
Grade	3.14* 3.14*	.22 2.16	2.50 2.22	.09 .46	29.69*** 25.14***	6.24** 8.86***
Grade x OM	1.23 1.23	.58 .44	1.25 1.25	.37 .18	1.88 1.46	.58 .22
Repeated Subjects Analyses						
Time	13.57*** 13.58***	62.69*** 84.12***	2.38 2.44	16.15*** 29.90***	.08 3.05	5.45* 3.10
OM x Time	.63 .63	1.46 1.30	.30 .38	.54 .40	.30 .34	.89 .47
Grade x Time	3.35* 3.35*	.96 .87	.76 .88	7.16*** 7.58***	27.29*** 31.68***	7.58*** 2.61
OM x Grade x Time	.58 .58	.62 .91	.93 .92	2.19 1.62	1.37 1.15	.88 .98

[a]All standard scores set at ≤ 160. [b]Pillai's criterion for statistical inference.
[c]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001.

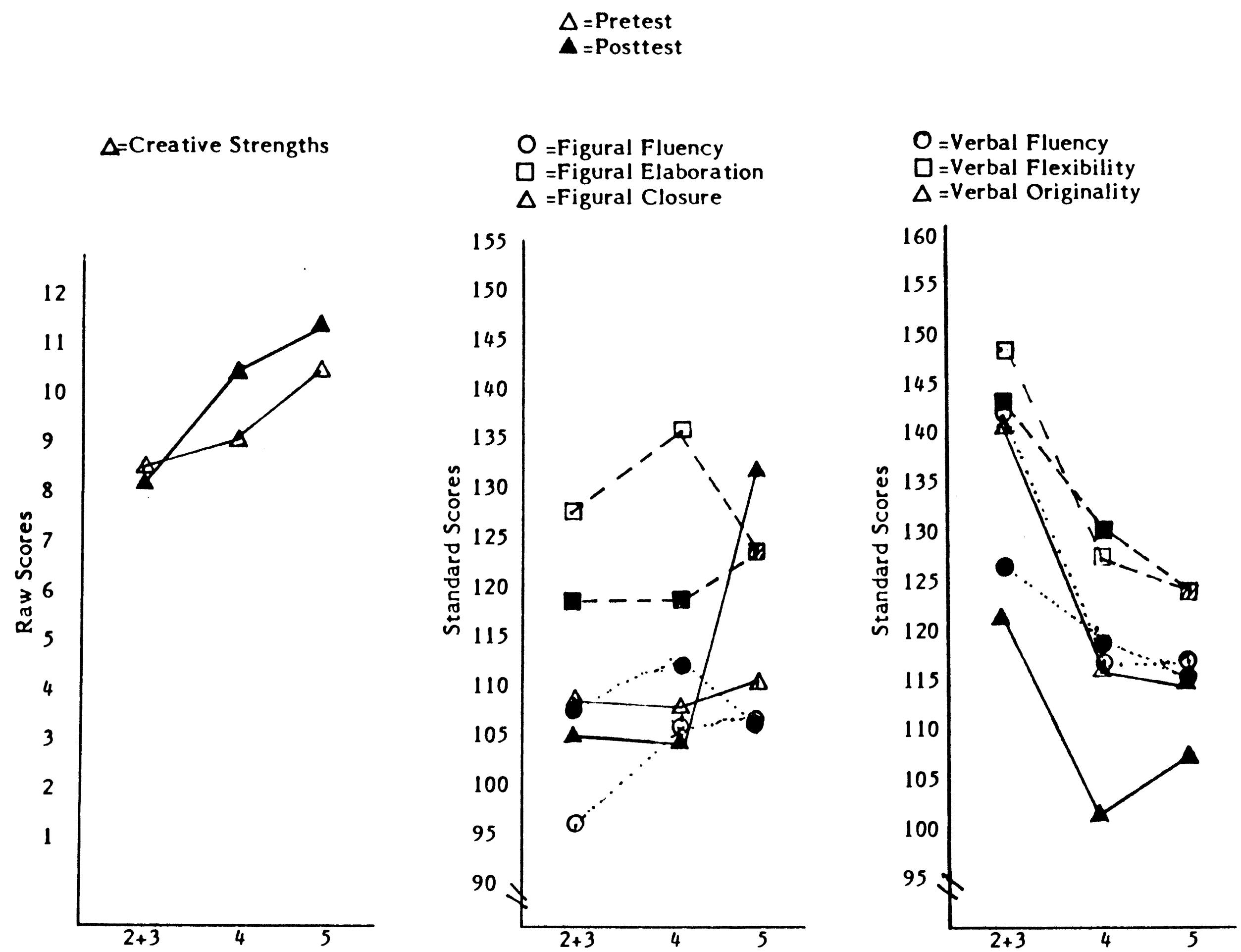


Figure P-11. Graph of the Interaction of Grade x Time on Creativity[a]

Note. All verbal and the five figural norm-referenced TTCT scores are expressed in standard scores. Similes and the total of creative strengths are expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at $\leq$ 160.

Table P-12

Intercorrelations of the Three Creative Dependent Variables[a]

	Similes	Mean Verbal TTCT	Figural TTCT Creativity Index
Similes	--		
Mean Verbal TTCT	-.13[b] -.08[c] -.07[d] .15[e]	--	
Figural TTCT Creativity Index	.33** .27** .28** .06	.12 .19* .18 .12	--

Note. The correlations are obtained from four different analyses.

[a]All standard scores set at $\leq$ 160.
[b]3x2, OM x Time mixed model.
[c]3x2x2, OM x Time x Experience in Enrichment Program.
[d]OM x Prior OM Experience x Time.
[e]MANCOVA.

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

Table P-13

Cell Means and Standard Deviations of Three Creative Dependent Variables, Odyssey of the Mind (OM) Effort, and Experience in the Enrichment Program[a]

Experience in Enrichment Program	OM Effort	*n*	Similes $\bar{X}$	Similes SD	Mean Verbal TTCT $\bar{X}$	Mean Verbal TTCT SD	Figural TTCT Creativity Index $\bar{X}$	Figural TTCT Creativity Index SD
Pretest Means								
< 13 Months	Non-OM	18	31.22	9.01	135.00	21.64	115.49	12.20
	OM-lo	11	48.18	11.17	128.94	15.74	122.87	9.77
	OM-hi	11	41.27	22.49	133.03	18.59	124.07	14.82
≥ 13 Months	Non-OM	33	41.58	14.44	128.83	20.25	121.52	12.88
	OM-lo	21	45.67	11.50	126.65	18.53	127.61	9.19
	OM-hi	22	54.73	19.53	123.80	16.60	128.95	9.96
Posttest Means								
< 13 Months	Non-OM	18	43.33	17.10	125.57	16.69	118.99	7.36
	OM-lo	11	51.09	13.03	117.79	16.45	126.47	13.58
	OM-hi	11	54.27	20.98	130.30	21.44	126.95	10.60
≥ 13 Months	Non-OM	33	50.03	15.38	117.56	14.82	127.16	10.53
	OM-lo	21	48.57	15.32	117.48	19.04	130.19	11.76
	OM-hi	22	64.55	16.08	124.55	17.46	134.96	11.64

Note. The verbal and figural TTCT scores are expressed in standard scores. Similes is expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at ≤ 160.

Table P-14

Summary of Mixed Model Creative MANOVA[a] for Odyssey of the Mind (OM) Effort and Experience in the Enrichment Program: Univariate Fs for Three Creative Dependent Variables[b]

Source	Univariate Fs		
	Dependent Variables		
	Similes	Mean Verbal TTCT	Figural TTCT Creativity Index
Between Subjects Analyses			
OM	9.27**	.63	7.96**
Experience in Enrichment Program	6.50**	3.26	11.83**
Experience x OM	2.40	.37	.22
Repeated Subjects Analyses			
Time	28.37**	21.26**	13.84**
OM x Time	2.29	3.78*	.29
Experience x Time	.60	.06	.39
Experience x OM x Time	.12	.23	.229

[a]No correlations between variables were $\geq$ .30 (see Table P-12, Appendix P), therefore a univariate interpretation was appropriate.
[b]All standard scores set at $\leq$ 160.

*p < .05. **p < .001.

Table P-15

Cell Means and Standard Deviations of Three Creative Dependent Variables, Odyssey of the Mind (OM) Effort, and Prior Experience on an OM Team[a]

			Dependent Variables					
Prior Experience on an OM Team	OM Effort	n	Similes $\bar{X}$	Similes SD	Mean Verbal TTCT $\bar{X}$	Mean Verbal TTCT SD	Figural TTCT Creativity Index $\bar{X}$	Figural TTCT Creativity Index SD
Pretest Means								
No Prior Experience	Non-OM	39	34.33	12.38	134.40	20.29	118.18	13.64
	OM-lo	12	45.58	12.04	128.89	15.99	125.40	11.23
	OM-hi	10	38.30	20.45	130.07	17.35	126.02	16.37
Experience on an OM Team	Non-OM	12	49.58	11.02	119.97	18.99	123.32	9.27
	OM-lo	20	47.10	11.06	126.57	18.54	126.33	8.62
	OM-hi	23	55.44	19.75	125.49	17.85	127.89	9.58
Posttest Means								
No Prior Experience	Non-OM	39	45.72	16.80	122.86	15.16	123.14	9 .28
	OM-lo	12	48.83	14.22	118.94	15.46	126.10	12.79
	OM-hi	10	54.50	19.28	129.00	21.56	125.58	11.36
Experience on an OM Team	Non-OM	12	54.00	12.47	112.33	15.83	127.97	12.67
	OM-lo	20	49.80	14.87	116.77	19.59	130.60	12.05
	OM-hi	23	64.00	17.36	125.36	17.79	135.20	10.94

Note. The verbal and figural TTCT scores are expressed in standard scores. Similes is expressed in raw scores. See Chapter III, Instrumentation.

[a]All standard scores set at $\leq$ 160.

Table P-16

Summary of Mixed Model Creative MANOVA[a] for Odyssey of the Mind (OM) Effort and Prior Experience on an OM Team: Univariate Fs for Three Creative Dependent Variables[b]

Source	Univariate Fs Dependent Variables Similes	 Mean Verbal TTCT	 Figural TTCT Creativity Index
Between Subjects Analyses			
OM	4.20*	.85	3.40*
Prior Experience on an OM Team	22.06**	3.96*	13.98**
Prior Experience x OM	1.91	1.02	.20
Repeated Subjects Analyses			
Time	28.97**	21.25**	14.15**
OM x Time	2.13	3.08*	.59
Prior Experience x Time	2.75	1.68	1.05
Prior Experience x OM x Time	.43	.11	.83

[a]No correlations between variables were $\geq$ .30 (see Table P-12, Appendix P), therefore a univariate interpretation was appropriate.
[b]All standard scores set at $\leq$ 160.

*p < .05. **p < .001.

Table P-17

Summary of the Creative Stepwise Multiple Regression Analyses: Contributions of 19 Predictor Variables to Regression Equations for Three Creative Dependent Variables[a]

Predictor Variables	Similes			Mean Verbal TTCT			Figural TTCT Creativity Index		
Statistics for Prediction Equation for all Predictor Variables									
Multiple R		.56			.62			.52	
R^2		.31*** [b]			.39***			.27***	
Adjusted R^2		.30			.37			.25	
F		25.36***			17.54***			13.93***	
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange
Pretest Variables									
Similes	.54	.51	41.07***				.26	.36	16.81***
Verbal Fluency				.44	.53	45.48***			
Verbal Flexibility									
Verbal Originality									
Figural Fluency							.18	.25	9.61**
Figural Originality									
Figural Titles									
Figural Elaboration							.18	.29	11.55***
Figural Closure									
Creative Strengths	1.29	.22	7.36**						
Posttest Variables									
Similes				.20	.56	6.54**			
Mean Verbal									
Figural Creativity Index									
Control Variables									
Months in program									
Years in OM									
Grade				8.81	.24	4.83*			
Teacher				8.23	.22	6.64**			
OM									
Sex									

Note. Three of the categorical variables, Grade, Teacher, and OM were coded into dummy variables to satisfy the assumptions required for regression analyses (Cohen & Cohen, 1983; Tabachnick & Fidell, 1983).

[a]All standard scores set at ≤ 160.
[b]The reported F ratios resulting from stepwise regression represent an inflated alpha Type I error. Greater confidence can be placed in the significance levels associated with Multiple R^2 (Tabachnick & Fidell, 1983).

*$p < .05$. **$p < .01$. ***$p < .001$.

Table P-18

Within Cells Correlations of Four Affective Dependent Measures Showing Relationship of Pretest[a], Posttest[b], and Adjusted Posttest Measures[c]

	Sears	I+	I-	Ideas
Sears				
I+	.37[a]** .12[b] .24[c]*			
I-	-.14 -.10 -.11	.16 .11 .12		
Ideas	.64** .26** .45**	.21* .12 .09	-.28** -.12 -.10	

[a]Pretest correlations are above posttest data.
[b]Posttest correlations are in the middle position.
[c]Adjusted posttest correlations are reported in the lowest position.

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

Table P-19

Cell Means and Standard Deviations of Affective Dependent Variables, Odyssey of the Mind (OM) Effort, and Experience in Enrichment Program

Experience in Enrichment Program	OM Effort	N	Sears $\overline{X}$	Sears SD	I+ $\overline{X}$	I+ SD	I- $\overline{X}$	I- SD	Ideas $\overline{X}$	Ideas SD
Pretest Means										
< 13 Months	Non-OM	17	186.53	31.45	14.29	2.09	8.82	3.97	34.18	5.79
	OM-lo	9	187.56	20.35	14.00	1.87	9.56	2.30	35.00	6.71
	OM-hi	11	183.46	27.85	12.82	2.64	9.46	2.66	36.82	5.00
≥ 13 Months	Non-OM	32	179.38	32.68	13.34	2.32	10.63	3.37	32.63	7.02
	OM-lo	20	178.10	34.15	13.70	2.36	8.85	2.54	35.05	7.10
	OM-hi	22	186.32	32.88	13.68	2.36	11.23	3.05	35.36	6.84
Posttest Means										
< 13 Months	Non-OM	17	190.82	20.85	14.29	2.54	9.59	3.83	35.82	4.20
	OM-lo	9	182.89	11.31	14.22	2.59	9.33	3.43	33.11	4.65
	OM-hi	11	172.46	41.31	14.18	1.89	11.09	2.12	36.64	6.42
≥ 13 Months	Non-OM	32	177.53	33.491	13.84	2.27	10.75	2.89	35.50	7.06
	OM-lo	20	163.05	38.79	12.90	3.26	10.45	2.61	35.30	6.25
	OM-hi	22	187.09	34.09	14.41	2.34	11.59	2.58	39.00	6.47

Table P-20

Summary of Mixed Model Affective MANOVA for Odyssey of the Mind (OM) Effort and Experience in the Enrichment Program: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Four Affective Dependent Variables

		Stepdown Fs/ Univariate Fs			
		Dependent Variables			
Source	Multivariate Global Fs	Sears	I+	I-	Ideas
Between Subjects Analyses					
OM	2.01*[a]	.71[b] .71	.03 .20	2.78 2.41	4.54* 2.45
Experience in Enrichment Program	1.79	.96 .96	.29 .73	4.00* 3.80	1.83 .00
Experience x OM	.77	1.27 1.27	.37 .96	.53 .47	.93 .32
Repeated Subjects Analyses					
Time	4.84***	2.78 2.78	2.36 1.80	3.85* 5.50*	9.45** 6.84**
OM x Time	1.42	2.19 2.19	2.42 2.84	.48 .55	.67 1.98
Experience x Time	.91	.20 .20	.21 .27	.09 .10	3.14 2.50
Experience x OM x Time	1.26	1.66 1.66	1.11 1.01	2.02 2.06	.29 .33

[a]Wilks' Lambda criterion for statistical inference.
[b]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001.

Table P-21

Cell Means and Standard Deviations of Affective Dependent Variables, Odyssey of the Mind (OM) Effort and Prior Experience in OM

Prior Experience on an OM Team	OM Effort	n	Sears $\overline{X}$	Sears SD	I+ $\overline{X}$	I+ SD	I- $\overline{X}$	I- SD	Ideas $\overline{X}$	Ideas SD
Pretest Means										
No Prior Experience	Non-OM	37	182.62	31.10	13.54	2.50	9.76	3.61	33.38	6.41
	OM-lo	11	190.09	19.28	14.00	1.73	9.09	2.34	35.27	6.51
	OM-hi	10	188.50	29.14	12.50	2.55	9.80	2.70	36.20	5.51
Experience on an OM Team	Non-OM	12	179.50	36.43	14.08	1.31	10.75	3.84	32.50	7.42
	OM-lo	18	175.50	35.03	13.67	2.47	9.06	2.58	34.89	7.24
	OM-hi	23	184.00	32.15	13.78	2.35	11.00	3.12	35.70	6.64
Posttest Means										
No Prior Experience	Non-OM	37	182.19	27.52	13.89	2.58	10.11	3.13	35.84	5.99
	OM-lo	11	180.55	11.52	14.18	2.48	9.00	3.23	34.09	4.83
	OM-hi	10	178.00	44.51	14.50	2.17	11.20	1.87	36.70	6.82
Experience on an OM Team	Non-OM	12	182.00	38.64	14.33	1.50	11.08	3.63	34.92	6.93
	OM-lo	18	162.28	40.91	12.78	3.35	10.78	2.49	34.94	6.45
	OM-hi	23	184.04	33.68	14.26	2.22	11.52	2.64	38.87	6.33

Table P-22

Summary of Mixed Model Affective MANOVA for Odyssey of the Mind (OM) Effort and Prior Experience on an OM Team: Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Four Affective Dependent Variables

Source	Multivariate Global Fs	Stepdown Fs/ Univariate Fs			
		Dependent Variables			
		Sears	I+	I-	Ideas
Between Subjects Analyses					
OM	1.54[a]	.71[b] .71	.08 .23	2.42 2.10	2.95 2.20
Prior Experience on an OM Team	1.93	.76 .76	.21 .01	2.28 2.88	4.37* .49
Prior Experience x OM	.54	.68 .68	.56 1.00	.01 .01	.92 .21
Repeated Subjects Analyses					
Time	4.76***	2.73 2.73	2.46 1.83	3.77 5.45*	9.17** 6.71*
OM x Time	1.60	2.38 2.38	3.04 3.31*	.24 .32	.83 2.10
Prior Experience x Time	.88	.00 .00	2.15 2.13	.92 .60	.48 .16
Prior Experience x OM x Time	.78	.57 .57	.88 .80	1.30 1.49	.41 .31

[a]Wilks' Lambda criterion for statistical inference.
[b]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001.

Table P-23

Summary of the Affective Stepwise Multiple Regression Analyses: Contributions of Predictor Variables to Regression Equations for Four Affective Dependent Variables[a]

	Affective Posttest Dependent Variables											
Predictor Variables		Sears			I+			I-			Ideas	
Statistics for Prediction Equation for all Predictor Variables												
Multiple R		.82			.63			.57			.67	
R^2		.67**[a]			.39			.32**			.45**	
Adjusted R^2		.65			.38			.31			.43	
F		34.76***			23.20***			25.85***			21.95***	
Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Pretest Variables												
Sears	.71	.66	91.88***									
I+				.37	.52	40.17***						
I-							.48	.53	42.23***			
Ideas										.19	.20	5.70*
Posttest Variables												
Sears				.02	.27	11.02***				.11	.58	53.30***
I+	1.88	.16	5.33*									
I-												
Ideas	1.85	.35	25.93***									
Control Variables												
Months in program										.09	.17	5.62*
Years in OM							.52	.17	4.83*			
Grade												
Teacher				1.60	.27	1.27						
OM	-13.00	-.17	2.36*									
Sex	-9.91	-.15	5.84*									

Note. Three of the control variables, Grade, Teacher, and OM were coded into dummy variables to satisfy the assumptions required for regression analyses (Cohen & Cohen, 1983; Tabachnick & Fidell, 1983).
Note. The predictor variables in the table with negative signs are suppressor variables, in actuality carrying very low correlation with the dependent variable of that equation. The suppressor variables (OM and Sex) suppress variance in other possible predictor variables which are irrelevant to the prediction of the dependent variable (Tabachnick & Fidell, 1983).

[a]The reported F ratios resulting from stepwise regression represent an inflated alpha Type I error. Greater confidence can be placed in the significance levels associated with Multiple R^2 (Tabachnick & Fidell, 1983).

*$p < .05$. **$p < .01$. ***$p < .001$.

Table P-24

Summary of Creative MANCOVA for Odyssey of the Mind (OM) Effort: Multiple Regression Statistics, Multivariate Global Fs, Stepdown Fs, and Univariate Fs for Ten Creative Variables[a]

		Stepdown Fs/Univariate Fs			
		Dependent Variables			
		Similes	Verbal TTCT		
Source	Multivariate Global F		Fluency	Flexi-bility	Origi-nality
Within Cells Regression Analysis of Covariates					
Regression F	2.90***[b]				
Multiple R		.57***[c]	.59	.62	.60
F-tests for Variables		3.14*** 4.04***	3.23*** 4.39***	3.55*** 5.25***	2.87** 4.78***
MANCOVA Between Subjects Analysis					
OM	1.31	2.90 3.40*	1.68 2.60	1.74 5.05	1.90 4.06*

	Stepdown Fs/Univariate Fs					
	Dependent Variables					
	Figural TTCT					
	Fluency	Origi-nality	Titles	Elabo-ration	Closure	Creative Strengths
Within Cells Regression Analysis of Covariates						
Regression F						
Multiple R	.51	.54	.49		.51	.53
F-tests for variables	2.30*** 2.30***	3.48*** 3.51***	2.66** 2.63**	2.92** 3.02***	5.14*** 5.36***	2.87** 3.36***
MANCOVA Between Subjects Analysis						
OM	.26 .26	.79 .44	.19 .18	.10 .11	1.03 .95	2.34 1.87

[a]All standard scores set at $\leq$ 160.
[b]Pillai's criterion for statistical inference.
[c]Stepdown Fs are above univariate Fs.

*p < .05. **p < .01. ***p < .001.

Table P-25

Intercorrelations of the Ten Creative and Four Affective Dependent Variables:

Correlations of Posttest Scores with Posttest and Pretest Scores[a]

	Creative Posttest Dependent Variables										Affective Posttest Dependent Variables			
	Similes	Verbal TTCT			Figural TTCT									
		Fluency	Flexibility	Originality	Fluency	Originality	Titles	Elaboration	Closure	Creative Strengths	Sears	I+	I-	Ideas
Creative Variables														
Similes	.53*** -	-.02 .11	-.06 .10	-.02 .04	.08 .02	.12 .11	.20* .08	.18* .27**	.30*** .16*	.33*** .36***	-.02 .15*	.21* .26**	.21* .15*	-.01 .18*
Verbal TTCT														
Fluency	-.14[b] -.11	.49*** --	.51*** .88***	.50*** .84***	.17* .30***	-.03 .13	.07 -.04	.11 .08	-.09 -.03	-.20** -.04	.10 .06	-.08 .08	.01 .13	-.09 .06
Flexibility	-.22** .10	.45*** .88**	.51*** --	.47*** .83***	.21** .25**	-.06 .11	.02 -.03	.09 .13	-.16* -.09	-.20** -.01	-.11 .04	-.02 .08	-.03 .14	-.10 .03
Originality	-.20 -.04	.47*** .84***	.49*** .83***	.45*** --	.16* .17*	-.06 .09	.05 -.04	.06 .12	-.15 .05	-.25** -.04	-.08 .07	-.03 .07	-.01 .13	-.08 .08
Figural TTCT														
Fluency	.12 .02	.25** .30***	.10 .25**	.09 .17*	.37*** --	.38*** .40***	-.09 -.05	.01 .14	.29*** -.03	.03 .08	.15 .06	-.07 .02	-.05 .07	.13 .29***
Originality	.21** .11	.19* .13	.09 .11	.14 .09	.20* .40***	.45*** --	-.06 -.06	.04 -.11	.23** .13	.11 .15	.25** .14	-.06 -.01	-.08 .01	.14 .19*
Titles	.15 .08	.04 -.04	.02 -.03	-.04 -.04	.10 -.05	.02 -.06	.32*** --	.01 .20*	.04 .10	.20* .36***	.12 .01	.05 -.07	.05 .03	.16* -.04
Elaboration	.27** .27**	.11 .08	.13 .13	.13 .12	.14 .14	.21** -.11	.04 .20*	.42*** --	-.01 .09	.31*** .52***	-.03 -.02	.09 .05	.25** .05	.10 .17*
Closure	.12 .16	-.02 -.03	-.01 -.09	.01 .05	-.03 -.03	.14 .13	-.17* .10	.08 .09	.18* --	.14 .26**	-.02 -.12	.01 -.10	.14 .07	.03 -.06
Creative Strengths	.33*** .36***	.04 -.04	-.01 .05	.11 -.04	-.11 .08	.19* .15	.13 .36***	.23** .52***	.26** .26**	.40*** --	.07 .03	-.05 -.03	.15* .09	.15 .16*
Affective Variables														
Sears	.14 .15*	.20* .06	.18 .04	.11 .07	.18* .06	.25** .14	-.14 .01	-.03 -.02	-.06 -.12	.02 .03	.69*** --	.32*** .37***	-.02 -.15	.42*** .59***
I+	.22** .26**	.05 .08	.09 .08	.05 .07	.17* .02	.07 -.01	-.08 -.07	.08 .05	-.01 -.10	-.04 -.05	.22** .37***	.52*** --	.08 .14	.08 .17*
I-	.23** .15*	-.01 .13	-.03 .14	-.07 .13	-.08 .07	-.02 .01	.09 .03	.03 .05	.15 .07	.01 .09	-.13 -.15	.13 .14	.53*** --	-.18* -.13
Ideas	.12 .29***	.10 .06	.11 .03	.10 .08	.21** .29***	.22** .19*	.04 -.04	.14 .17*	-.01 -.06	.16* .16*	.42*** .59***	.12 .17*	-.09 -.13	.44*** --

Note. This table reflects all possible pre- and posttest comparisons for each variable. Therefore, mirror image values should not be expected on either side of the diagonal. Example: Sears pre to I+ post - .32; Sears post to I+ pre - .22; Sears post to I+ post = .37.
Note. n ranged from 108 to 120 for the various comparisons.

[a] All standard scores set at ≤ 160.
[b] Correlations with pretest dependent variable appear above correlations with posttest dependent variable.

*p < .05. **p < .01. ***p < .001.

Table P-26

Summary of the Affective Stepwise Multiple Regression Analyses: Contributions of 20 Creative and Eight Affective Predictor Variables to Regression Equations for Four Affective Dependent Variables[a]

Predictor Variables	Affective Posttest Dependent Variables											
	Sears			I+			I-			Ideas		
Statistics for Prediction Equation for all Predictor Variables												
Multiple R		.80			.64			.60			.67	
R^2		.64***[b]			.41***			.36***			.44***	
Adjusted R^2		.63			.39			.34			.43	
F		46.72***[b]			18.40***			19.73***			28.10***	
Contribution of Predictor Variables	B	β	F ch	B	β	F ch	B	β	F ch	B	β	F ch
Pretest Variables												
Creative Construct												
Similes												
Verbal Fluency												
Verbal Flexibility												
Verbal Originality												
Figural Fluency												
Figural Originality				-.03	-.24	6.31**	-.03	-.17	4.60*			
Figural Titles												
Figural Elaboration							.04	.25	6.70**			
Figural Closure												
Creative Strengths												
Affective Construct												
Sears	.56	.52	105.09***[b]									
I+				.47	.42	39.31***						
I-							.50	.54	43.80***			
Ideas												
Posttest Variables												
Creative Construct												
Similes				.03	.20	6.72***						
Verbal Fluency												
Verbal Flexibility												
Verbal Originality												
Figural Fluency	-.29	-.14	5.75*							.08	.21	9.95**
Figural Originality												
Figural Titles												
Figural Elaboration										.05	.17	5.15*
Figural Closure												
Creative Strengths												
Affective Construct												
Sears				.02	.30	11.36***				.11	.59	61.16***
I+	1.80	.14	4.75*									
I-												
Ideas	2.08	.39	29.81***									

Note. The predictor variables in the table with negative signs are suppressor variables, in actuality carrying very low correlation with the dependent variables in that equation. The suppressor variables suppress variance in other possible predictor variables which are irrelevant to the prediction of the dependent variable (Tabachnick & Fidell, 1983). That is, figural originality is a suppressor variables for I+ and I-, as is figural fluency for Sears.

[a] All standard scores set at ≤ 160.
[b] The reported F ratios resulting from stepwise regression represent an inflated alpha Type I error. Greater confidence can be placed in the significance levels associated with Multiple R^2 (Tabachnick & Fidell, 1983).

*$p < .05$. **$p < .01$. ***$p < .001$.

Table P-27

Summary of the Creative Stepwise Multiple Regression Analyses: Contributions of 20 Creative Predictor Variables to Regression Equations for Ten Creative Dependent Variables[a]

Predictor Variables	Creative Posttest Dependent Variables											
	Similes			Verbal TTCT								
				Fluency			Flexibility			Originality		
Statistics for Prediction Equation for all Predictor Variables												
Multiple R		.56			.91			.90			.89	
R^2		.31** [b]			.83**			.82**			.79**	
Adjusted R^2		.30			.83			.81			.78	
F		25.36***			186.84***			123.20***			67.95***	
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Pretest Variables												
Similes	.54	.51	41.07***									
Verbal Fluency												
Verbal Flexibility							.57	.58	5.55*			
Verbal Originality												
Figural Fluency				.19	.17	18.20***	-.11	-.10	5.00*	-.12	-.12	5.41*
Figural Originality										.15	.17	5.80*
Figural Titles												
Figural Elaboration												
Figural Closure												
Creative Strengths	1.29	.22	7.36**									
Posttest Variables												
Similes										-.11	-.11	5.97*
Verbal Fluency							.86	.88	374.78***	.73	.84	269.88***
Verbal Flexibility				.89	.88	374.78***				.37	.42	17.71***
Verbal Originality				.41	.36	23.14***	.37	.33	17.71***			
Figural Fluency												
Figural Originality												
Figural Titles												
Figural Elaboration												
Figural Closure										.11	.11	5.05*
Creative Strengths												

(table continues)

Table P-27 (continued)

Predictor Variables	Creative Posttest Dependent Variables: Figural TTCT: Fluency			Originality			Titles			Elaboration			Closure			Creative Strengths		
Statistics for Prediction Equation for all Predictor Variables																		
Multiple R	.60			.58			.56			.65			.44			.71		
R^2	.36**			.34**			.31**			.42**			.20**			.51**		
Adjusted R^2	.32			.32			.28			.40			.17			.48		
F	10.01***			18.88***			9.97***			20.27***			9.08***			22.57***		
Contribution of Predictor Variables	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange	B	β	Fchange
Pretest Variables																		
Similes													.31	.30	11.57***			
Verbal Fluency							.17	.20	5.63*	.18	.18	5.99*						
Verbal Flexibility				-.21	-.18	5.36*												
Verbal Originality																-.04	-.26	13.06***
Figural Fluency	.19	.19	4.65*										.28	.27	10.18**			
Figural Originality	-.26	-.32	5.74*	.48	.46	31.16***												
Figural Titles							.27	.25	8.62**									
Figural Elaboration										.30	.29	12.80						
Figural Closure							-.37	-.24	8.40**									
Creative Strengths	-.99	-.18	4.55*													.32	.29	14.18***
Posttest Variables																		
Similes																		
Verbal Fluency	.23	.27	10.90***															
Verbal Flexibility							-.20	-.20	4.67*									
Verbal Originality																		
Figural Fluency				.39	.31	14.53***												
Figural Originality	.31	.39	20.92***							-.24	-.24	10.16**				.03	.17	6.10*
Figural Titles																.05	.27	14.37***
Figural Elaboration	.14	.18	5.21*													.08	.51	40.15***
Figural Closure																		
Creative Strengths							2.09	.36	16.71***	3.32	.51	40.15***	.95	.18	3.98*			

Note. The predictor variables in the table with negative signs are suppressor variables, in actuality carrying very low correlation with the dependent variables in that equation. The suppressor variables suppress variance in other possible predictor variables which are irrelevant to the prediction of the dependent variable (Tabachnick & Fidell, 1983). For example, figural originality is a suppressor variable in the equation for elaboration.

[a]All standard scores set at ≤ 160.
[b]The reported F ratios resulting from stepwise regression represent an inflated alpha Type I error. Greater confidence can be placed in the significance levels associated with Multiple R^2 (Tabachnick & Fidell, 1983).

*p < .05. **p < .01. ***p < .001.

VITA

Anne Sonya Fishkin

Candidate for the Degree of

Doctor of Philosophy

Thesis: EFFORT OF ODYSSEY OF THE MIND CREATIVE PROBLEM SOLVING TEAMS: EFFECTS ON CREATIVITY, CREATIVE SELF-CONCEPT, LOCUS OF CONTROL AND GENERAL SELF-CONCEPT IN GIFTED CHILDREN

Major Field: Applied Behavioral Studies

Specialization: Special Education

Biographical:

Personal Data: Born in Brooklyn, New York, November 25, 1938 to Harry A. and Judith E. Pollack. Two daughters--Ruth Ellen Fishkin Campbell and Amy Lynne Fishkin.

Education: Early entrant student to college following two years at Forest Hills High School, Queens, New York; received Bachelor of Arts degree (liberal arts) from Shimer College, Mt. Carroll, Illinois in 1958; received Master of Arts degree (major, rehabilitation counseling; minor, experimental psychology) from the University of Colorado, Boulder, Colorado, 1970; completed requirements for the Oklahoma Standard Certification (school psychology) from Central State University, Edmond, Oklahoma, 1978; completed requirements for the Oklahoma Standard Certificate (learning disabilities) from Oklahoma State University, December, 1988; and completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1989.

Professional Experience: Research Assistant, Department of Counseling and Testing, 1963-64; Institute of Behavioral Sciences, 1963-66, University of Colorado, Boulder; Intern in School Psychology at Oklahoma City Public Schools, Fall, 1974, and at the Psychology Clinic, Central State University, Edmond, Oklahoma, Summer, 1975; Substitute teacher, Oklahoma City Public Schools, Spring, 1975; Enrichment Teacher with gifted students, Grades 2-12 in Guthrie Public Schools, 1975-76, and with Grades 1-6 in Millwood Elementary School, 1976-78; Enrichment Teacher, Edmond Public Schools,