A PRELIMINARY INVESTIGATION INTO
THE EFFECTS OF TOKEN REINFORCEMENT
ON ONE ASPECT OF CREATIVITY:
AS MEASURED BY THE WALLACH-KOGAN CREATIVITY TEST

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Ferris Owen Henson II, B.S., M.A.

*****

The Ohio State University
1974

Reading Committee: Approved by
Thomas M. Stephens, Ph.D.
John O. Cooper, Ph.D.
Henry Leland, Ph. D.

[Signature]
Adviser
Department of Exceptional Children
To the often times deaf,
To the often times blind,
To those we love,
To those we bind......
Ourselves together for eternity.

for Karen
ACKNOWLEDGEMENTS

First, I would like to express my appreciation to my advisor, Thomas Stephens, for his advice and direction in the present study. Secondly, I would like to thank the children who served as subjects for the study, for without their help this investigation would not have been possible.

I would also like to thank John Cooper for his suggestions on the design used, as well as his constant interest in the incoming data. Additionally, I would like to thank Henry Leland for his advice on the population sample used, and his time put forth reading the drafts and his suggestions on the notion of divergent responding.

Also, my appreciation is expressed to Calvin Anderson for his guidance throughout my program.

Finally, I would like to thank my daughter, Celeste, and dearest friend, Leatha, for their perseverance, as well as the partial schedule of verbal reinforcement they provided me in maintaining my writing behavior.

Special thanks goes to the memory of my wife, Karen, and her spirit which helped me through my times of stress.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF GRAPHS</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I.  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1. Background</td>
<td>1</td>
</tr>
<tr>
<td>2. Rationale</td>
<td>3</td>
</tr>
<tr>
<td>3. Problem Statement</td>
<td>5</td>
</tr>
<tr>
<td>4. Operational Definitions</td>
<td>7</td>
</tr>
<tr>
<td>II. REVIEW OF THE RESEARCH</td>
<td>9</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2. Definitions of Creativity</td>
<td>10</td>
</tr>
<tr>
<td>3. Creativity as a Process</td>
<td>13</td>
</tr>
<tr>
<td>4. Creativity as Personality</td>
<td>18</td>
</tr>
<tr>
<td>5. Creativity as a Product</td>
<td>21</td>
</tr>
<tr>
<td>6. Research Most Related to</td>
<td></td>
</tr>
<tr>
<td>Present Investigation</td>
<td>23</td>
</tr>
<tr>
<td>7. Summary</td>
<td>35</td>
</tr>
<tr>
<td>III. METHOD</td>
<td>36</td>
</tr>
<tr>
<td>1. Subjects</td>
<td>36</td>
</tr>
<tr>
<td>2. Setting</td>
<td>36</td>
</tr>
<tr>
<td>3. Instrument</td>
<td>37</td>
</tr>
<tr>
<td>4. Design</td>
<td>39</td>
</tr>
<tr>
<td>5. Baseline I</td>
<td>39</td>
</tr>
<tr>
<td>6. Treatment</td>
<td>41</td>
</tr>
<tr>
<td>7. Tokens and Exchange Period</td>
<td>42</td>
</tr>
<tr>
<td>8. Reintroduction of Reinforcement</td>
<td>43</td>
</tr>
</tbody>
</table>
IV. RESULTS ........................................ 44
   Gifted Group .................................. 44
   Average Group ................................. 45
   Learning Disabled Group ................. 50
   Statistical Analysis ....................... 52
   Summary ................................... 53

V. DISCUSSION, LIMITATIONS,
   IMPLICATIONS AND SUMMARY .......... 54
   Discussion ................................. 54
   Limitations ................................. 57
   Implications ............................... 59
   Recommendations for Further
   Research .................................. 60
   Summary .................................. 61

REFERENCES .................................................. 65
APPENDIX A .................................................. 70
APPENDIX B .................................................. 75
APPENDIX C .................................................. 78
APPENDIX D .................................................. 79
APPENDIX E .................................................. 80
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Summary of Mean Scores across the Varying Experimental Conditions</td>
<td>46</td>
</tr>
<tr>
<td>2. Summary Table of Kruskal-Wallis One Way Analysis of Variance by Ranks</td>
<td>52</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Graphic Representation of Design</td>
<td>40</td>
</tr>
</tbody>
</table>
LIST OF GRAPHS

<table>
<thead>
<tr>
<th>Graph</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Graphic Representation of Data for</td>
<td></td>
</tr>
<tr>
<td>Subjects $G_1$ and $G_2$</td>
<td>47</td>
</tr>
<tr>
<td>II. Graphic Representation of Data for</td>
<td></td>
</tr>
<tr>
<td>Subjects $A_1$ and $A_2$</td>
<td>49</td>
</tr>
<tr>
<td>III. Graphic Representation of Data for</td>
<td></td>
</tr>
<tr>
<td>Subjects $LD_1$ and $LD_2$</td>
<td>51</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Background

Ever since the concept of creativity has become a relatively current issue in psychology and education (Guilford, 1950), problems concerning its definitions have arisen. The problem of definition probably stems from the particular educational/psychological approach used to investigate creativity. That is, in the past researchers have approached the topic from many viewpoints.

Bloomberg (1973) listed the following approaches: a) psychoanalytic, b) humanistic, c) environmental, d) associative, e) factorial, f) holistic, and g) cognitive-developmental. Other researchers (Rhodes, 1961; Taylor, 1959; and Goetz and Baer, 1973) have made reference to the ambiguities in definitions of creativity. While an obvious solution would be for researchers to agree on a definition of creativity, this agreement is unlikely to occur. Due to the ambiguity and divergency of definitions of creativity, this writer has not used the term "creativity" throughout the present investigation (except where necessary to convey the findings of other researchers). Instead, an approach investigating the maximizing of divergent responses was used.

An approach presented in three investigations (Goetz and Baer, 1973; Maloney and Hopkins, 1973; and Goetz and Salmonson, 1972) utilized diverse response patterns as dimensions of creativity (e.g., easel painting and block-building behavior). These variables are diverse responses emitted to a stimulus. These responses are de-
fined by the researcher in such a manner so as to facilitate accurate quantification.

Such a measurable variable, i.e., divergent responding, was of particular interest to the present investigation. Divergent responding was considered as the number of verbal responses emitted by the subject to a stimulus item from the Wallach-Kogan Creativity Test (Wallach and Kogan, 1965). Since reinforcement contingencies have been used with success in increasing a variety of response patterns (Hall, 1968; O'Leary, Becker, Evans, and Saudargas, 1969), it would be logical to approach divergent response patterns in a similar manner. One group of responses not often used in behavioral research is divergent responding to a given stimulus. The works of Goetz and Baer (1973), Maloney and Hopkins (1973), and Goetz and Salmonson (1972) were concerned with the effects of reinforcement on certain diverse response patterns. The results of all three of these studies showed an increase in the diversity of response patterns when reinforcement was introduced contingent upon certain operationally defined behaviors.

Out of the above information grew a need for further investigation into the effects of reinforcement on divergent responding, specifically, the number of verbal responses emitted to a verbal stimulus. A secondary need of such an investigation was to test for statistically significant differences in effects of treatment across different categories of children, i.e., average children, gifted children, and learning disabled children, in hopes of establishing a beginning empirical basis for application of the findings to the classroom setting. This would also allow for greater generality of effects. The present investigation sought to accomplish this twofold task.
In summary, the purpose of the present study was to investigate the effects of reinforcement upon divergent responding to items from the Wallach-Kogan Creativity Test. Additionally, a secondary purpose was to investigate for any statistically significant differences in effects across differing categories of children, i.e., average children, gifted children, and learning disabled children.

**Rationale**

While traditional research designs use a large number of subjects, the design employed in the present study used a small number of subjects (six in number). Additionally, studies involving only a small number of subjects use each subject as his own control rather than using a control group per se. This distinction, as well as others, is discussed below as a general rationale for using the design.

A misconception that behavioral research implies that the problem of treatment generalization to the population is determined by investigating only a small number of subjects was expressed by Kazdin (1973) in writing "Also, various elements of the experiment may delimit the generalization of the results (p. 528)."

Risely (1969) stated the following regarding inter-subject generality:

**Generality of effect across subjects is important,** of course, but inter-subject generality should be considered apart from the magnitude of the change. In behavior modification research, the usual approach is to work first with a few subjects intensively in order to discover procedures which produce significant changes in their behaviors, and only then to apply those procedures to other
subjects. By this process, inter-subject
generality is determined not simply as
a statement of the range of effects of a
specific procedure across subjects, but
rather as a description of the range of
variations of the procedure required to
produce significant changes in all sub-
jects (p. 11).

A concern for treatment variation across subjects is implicit for
applied research. "Significant change" refers to a change in topo-
graphy and/or rate of response to a more acceptable level (Risely,
1969). Statistical significance does not relate to this type of signi-
ficance.

It must be re-emphasized, in regard to inter-subject generality,
that single organism designs do not purport to be generalizable to
the population under study. Also, single organism designs allow for
needed change in intervention as individual subjects differ. The single
organism design allows for a flexible model of treatment which can
be modified with each subject rather than indicating the range of
effects of a single treatment across all subjects. In determining the
generality of effects of treatment across subjects, replication is
necessary.

Replication of an experiment can serve two functions in behavior
modification research; first, in an intra-subject sense to determine
functional relationships between variables, and second, in an inter-
subject sense to determine generality of treatment (Sidman, 1960).

An example of an indication of functional relationships between
variables is shown in the following hypothetical example. If the data
indicated that nine times out of nine times the introduction of a parti-
cular variable coincided contiguously with a slight increase in a par-
ticular response, it would provide a basis for describing the function-
al relationships between the variables under investigation.

One might note, however, that while this type of intra-subject replication is indicative of functional relationships, nothing is offered empirically concerning the generality of treatment. This is only accomplished by inter-subject replication.

If, after the data supported the notion that a functional relationship existed between two variables, the researcher were concerned with generality of treatment, he must repeat the same experiment on other subjects of the same species (Sidman, 1960). If inter-subject replications are successful, the treatment can be determined as generalizable to the population under study. Of course, the more times the study is replicated the more powerful will be the statements concerning generality of treatment.

**Problem Statement**

The hypotheses and questions raised in the present investigation were similar to those used by Ward, Kogan, and Pankove (1972) in studying incentive effects on children's creativity. Of major emphasis to the present study was to determine the effects of a reinforcement system (i.e., a token reinforcement system) upon the number of responses given to items on the Wallach-Kogan Creativity Test. A secondary concern was to determine if the effects of treatment differed significantly (statistically) across the differing categories or type of subjects used. Considering that the results of the Ward, et al., (1972) study were in conflict with other studies using reinforcement contingencies and divergent response patterns (Goetz and Baer, 1973; Goetz and Salmonson, 1972; and Maloney and Hopkins, 1973), a similar study was proposed, changing the design to a single
organism design.

Goetz and Baer (1973), Maloney and Hopkins (1973), and Goetz and Salmonson (1972), supported the hypothesis that reinforcement can increase divergent behaviors predefined by the experimenter. If such behaviors as diversity in block-building and increasing the use of parts of speech in story writing can be increased using reinforcement contingencies, then responses to items on the Wallach-Kogan Creativity Test may also be increased using these contingencies. Considering the success in using reinforcement to increase divergent response patterns the following two research questions were formulated.

Question I:

Can test scores on the Wallach-Kogan Creativity Test be positively affected by the use of reinforcement contingencies placed on the responses to the items of that test?

Question II:

If positive effects occur using reinforcement, do these effects vary across differing categories of children, specifically, the gifted child, the learning disabled child, and the average child?

Since Question II was treated statistically, the null hypothesis was formulated as follows:

\[ H_0: \text{There will be no statistically significant differences in mean gain scores, or effects of treatment, among the differing categories of subjects, i.e., the gifted child, the average child, and the learning disabled child.} \]

\[ H_0 = \bar{X}_a = \bar{X}_b = \bar{X}_c \]
Operational Definitions of Terms

For the purposes of this study the following operational definitions were used:

Divergent Responding, or Fluency—the number of verbal responses emitted to verbal stimuli. Appropriateness of the responses was determined by the experimenter using outside judges for inter-judge reliability ratings. Appropriateness was defined as the response fitting the context of the stimulus item.

Divergent Response Scores—number of verbal responses emitted to each item on the three verbal subtests of the Wallach-Kogan Creativity Test.

Average Children—children placed in a regular classroom and having a recorded standard I.Q. score, for purposes of categorization, between 85 and 115. Teacher and principal recommendations were also considered as were personal school records.*

Gifted Children—children either in a regular classroom, or in a special class for the gifted, and having a recorded standard I.Q. score, for purposes of categorization, of 140 or higher. Teacher and principal recommendations were also considered as were personal school records.*

Learning Disabled Children—children placed in a special class for the learning disabled, or in a regular classroom, and functioning significantly below grade placement or ability, as indicated by the child's school achievement record, the child's principal, and the child's teacher. For purposes of categorization, the child must also have a recorded standard I.Q. score of 80 or higher.

*Permission to review school records was obtained from the school's principal and the children's parents.
Reinforcement—the use of tokens contingent upon responding appropriately to the given verbal stimulus. Verbal praise was also presented with the token reinforcement. The tokens could then be traded for tangible items listed on a reinforcement list. (See Appendix E).
CHAPTER II

REVIEW OF THE RESEARCH

Introduction

The literature related to creativity contains several distinct, though not unrelated, categories of investigation. Three themes seem dominant. The first deals with creativity as a psychological process. The literature by Guilford (1950, 1956, 1959), Kubie (1958, 1960), Barron (1958), Maslow (1959), Hart (1950), Torrance (1960, 1965), Eisner (1963), and Fromm (1963), is apparently concerned with the psychological process of creativity. The process of creativity may be further subdivided into processes concerned with some form of learning theory, e.g., Guilford and Torrance, and processes as defined from a psychoanalytical point of view, e.g., Kubie and Fromm.

The second category of research deals with creativity in terms of personality traits and individual, or group characteristics. Some of this work is historical research and relies heavily on personal anecdotes and bibliographic studies of creative people in an experimental setting, involving test measures concerned with personality traits, e.g., dependence and independence, and conformity and non-conformity. Some research on the creative personality was done by Cattell and Drevdahl (1955, 1958), Getzels and Jackson (1962), Kogan and Wallach (1964), Anderson and Cropely (1966), MacKinnon (1962), and Cropely (1967).

The third area of inquiry into research on creativity is reflected in the works of Gordon (1961), Goetz and Salmonson (1972), Maloney
and Hopkins (1973), Ward, Kogan, and Pankove (1972), and Goetz and Baer (1973). While their work is generally classed according to output or product, specifically, the works of Goetz and Baer, Maloney and Hopkins, and Goetz and Salmonson, used an experimental analysis of behavior approach to creativity (i.e., divergent response patterns).

Much overlap in each of these areas exists. Of major concern, however, is the investigation into the functional relationships of the variables involved in certain aspects of creativity. After discussing briefly the diversity of opinions on the definition of creativity, a review of some of the research in each of the above mentioned areas is given, i.e., creativity as a process, creativity as personality, and creativity as a product. Finally, some comments are made concerning a recommended direction for research into the area of creativity, or divergent response patterns.

**Definitions of Creativity**

Creativity in its narrowest sense refers to the traits or abilities characteristic of creative people (Guilford, 1950). Guilford stated, "A creative pattern is manifest in creative behavior, which includes such activities as inventing, designing, contriving, composing, and planning (p. 444)."

Rhodes (1961) wrote of creativity as involving person, process, product, and press, press referring to the environment. Uniting all of these characteristics, he defined creativity as, "a noun naming the phenomena in which a person communicates a new concept (which is the product) (p. 305)."

Another approach to defining creativity has to do with its relation-
ship to intelligence. In this case, to ascribe the term "correct" and "good" to a response or response pattern is helpful (Jackson and Messick, 1965). In comparing or contrasting intelligence to creativity, one might denote an intelligent response as correct in terms of the response satisfying objective criteria, operating within the confines of logic, and relating to reality. The response can be determined as right or wrong. These types of criteria are applicable to operational definitions of intelligence as well as creativity in terms of test scores. The "goodness" of a response is more elusive, and hence more compliant to subjective criteria, and it may or may not fall within the realm of logic. There is, however, some criteria of appropriateness used in judging the response as "good" or creative. A quote from Santayana exemplifies this, "...for the existence of good in any form it is not merely consciousness but emotional consciousness that is needed. Observation will not do, appreciation is required, (Santayana, 1896, p. 16)."

Taylor (1959) followed a similar pattern in defining creativity. He proposed that creativity falls outside the realm of logic, or into the subjective realm. Creative thought, according to Taylor, is constrained by the rules of logic and the scientific method. After analyzing the diversity of definitions of creativity, Taylor constructed five levels of creative thought. First was expressive creativity which is simply independent expression. Next was inventive creativity which is when ingenuity is displayed. Inventive creativity does not contribute new ideas, but new uses for old ideas. The next stage is innovative creativity and this only emerges when one has obtained a fundamental knowledge of a given area, and can conceptualize to the point of modifying that particular area. The highest form of
creativity, according to Taylor, is emergentive creativity, which is conceptualizing an entirely new principle. Note that the logic involved in this framework is theoretical and the pragmatic problem of empirical observation, or measurement remains.

While these types of definitions do indeed enable one to grasp a broad notion of the concept of creativity, operationally they leave science little to investigate empirically. The reason is that many of these definitions can be classed into the realm of philosophy and theory, which, while not antagonistic to science in general, specifically fail to fulfill the requirement of quantification. How can one measure the goodness, richness, or appreciative qualities of a response or product from a scientific viewpoint?

A small group of researchers, namely, Goetz and Salmonson (1972), Maloney and Hopkins (1973), and Goetz and Baer (1973), seeing the need for the role of functional relationships between variables, have dealt with creativity (diverse response patterns) at this level. While the qualitative aspects previously discussed are not included in this approach, it is certainly plausible to proceed to work on research in creativity, but only at a level where quantification can occur. The necessity of sorting out the variables with which science can work, and obtaining data on the operations of these variables within the context of a given environment becomes paramount. In so doing, these researchers have dealt with the novel occurrence of a predefined behavior. Novel occurrence was defined as a given behavior never before displayed by the subject during the duration of the experiment. While this seems to be an oversimplification of the concept of creativity, it does allow for the term divergent response pattern to be investigated in such a manner so that functional
relationships between some of the variables can be examined. After
more is learned about these relationships, other criteria and varia-
bles of creativity may be empirically examined.

In summary, three areas of creativity research have been identi-
fied, namely, creativity as a process, creativity as personality, and
creativity as a product. Some of the problems involving definitions
of creativity have been posed, as well as a possible alternative which
can be taken. The review that follows is divided into these three
areas, but the major focus is on the last category, or creativity as
a product. In that category a more in depth analysis occurs.

Creativity as a Process

Creativity as a process deals with mediation and its concomitant
behavior. Mediation is not observable directly, and can only be
inferred from external behavior patterns or performance.

Guilford (1956), applying the principles of factorial design, wrote
that once one hypothesizes a specific unitary factor of the intellect, a
psychological test may be used to determine the nature and existence
of that factor. Guilford classified the intellect into two categories,
thinking and memory. The category of thinking is most related to
creativity as a process. Thinking is further delineated as, a) cogni-
tive factors, b) production factors, and c) evaluation factors, with
production factors being further classed as convergent and divergent
subfactors. Convergent thinking is that process measured by stan-
dardized intelligence tests, and divergent thinking is that related to
creativity measures.

In hypothesizing divergent thinking factors which require investi-
gation, Guilford referred to the following factors: synthesizing and
analyzing abilities; sensitivity to problems; novel ideas; flexibility (ease of changing mental set); re-organization, or re-defining of organized wholes; and the amount of complexity of conceptual structure with which an individual is comfortable in dealing. If these correlates of creativity can be supported, then it will be possible to devise a method of teaching creativity (Guilford, 1950).

Barron (1955) maintained that the creative act is disembodied from the process by limiting the inquiry to the creator's mental content at the moment of insight rather than seeking to understand the creator's highly organized system of responding. "There is good reason for believing, however, that originality is almost habitual with persons who produce a really singular insight, (Barron, 1955, p. 478)." According to Barron, voluminous productivity is the rule rather than the exception.

Using eight measures of originality (three by Guilford, the Rorschach Test, the Thematic Apperception Test, an anagram, and two measures constructed by Barron), the following hypotheses were confirmed, according to Barron (1955): a) creative persons prefer complexity and imbalance in phenomena; b) creative persons are more complex psychodynamically, with greater personal scope; c) creative persons are more independent in judgment; d) creative persons are more self-assertive and dominant; and e) creative persons reject suppression as a mechanism for control of impulses.

One should note that while these differences were statistically significant using a t test, not only are the conclusions hypothetical in nature, but also many of the terms are vague, e.g., personal scope. To define personal scope as a score on a test that measures personal scope is tautological. Likewise, to assess creativity using projec-
tive tests, in terms of originality, and then to report the results in a conclusive manner, ignores the subjectivity involved in interpreting such tests.

Another inquiry related to creativity as a process is an investigation by Torrance (1965). He hypothesized that there are five principles which, if followed by teachers, would enhance creativity in children. These may be summarized as follows: a) be respectful to unusual questions. That is, to accept a child's question as indicative of his interest in inquiry. b) Be respectful to unusual and imaginative ideas. This refers to responding to unusual ideas in a positive fashion so as to build up confidence in the child to continue thinking in imaginative ways. c) Show your pupils that their ideas have value. Borkan (1960) was cited as pointing out that kindergarten children need teacher assurance that their ideas have value. (Upon examining that source, this writer found that this statement by Borkan was not documented empirically. While one could not doubt the obvious implications of teacher assurance on children's ideas, one could doubt the credibility of such a statement when presented as factual when in fact no empirical evidence was cited to support the statement.) d) Have students practice without threat of evaluation. External evaluation is threatening and enhances defensiveness. e) Tie in evaluation with causes and consequences. Show the child that certain divergent behaviors have desirable consequences and others do not.

To test the above mentioned principles Torrance (1965) asked teachers to merely carry out these principles. Not much experimental control was used, and the teachers were simply asked to be as systematic as possible while using the principles. Data was collected using a questionnaire.
The results obtained from a study of this nature are relatively unspecific. The author stated that many teachers understood the principles and applied them with favorable results (subjectively determined by the teacher). On the other hand, other teachers tried to apply the principles but, according to Torrance (1965) did not fully understand their implication or meaning. That is quite understandable, as the principles were conveyed in vague terminology, e.g., be respectful to unusual questions. Other intervening variables which led to undesirable results were discussed by Torrance. These were; rigidity on the part of the teachers, preoccupation with time factors, authoritarian roles, defensiveness, insensitivity to students' emotional and intellectual needs, lacking intellectual energy, preoccupation with disciplinary matters, etc. While slight explanations of these variables were presented, it did little to enhance specificity or quantifiability of these variables.

Other researchers have theorized the process of creativity from a psychoanalytical point of view. While much can be written from this point of view, only the works of Kubie (1961), and Fromm (1963) are examined. Such works explicitly point out an example of elaborate hypothetical constructs.

Kubie (1961) stated that mental activity occurs in three operant symbol making systems. These are the unconscious system, the preconscious system, and the conscious system. Thinking occurs in the preconscious system, communication in the conscious system, and repression in the unconscious system. Creative thinking occurs as a result of an interaction between the preconscious system and the conscious system. Also, creative thinking is adversely affected by the unconscious system.
The preconscious system, according to Kubie (1961), is considered to be analogous to a data bank of information which can be brought to awareness at will. He pointed out that this system is differentiated from the other two in the following manner. The preconscious system is not the same as the unconscious system because the information from the preconscious can be retrieved at will while data from the unconscious cannot be brought to awareness volitionally. On the other hand, the preconscious system is differentiated from the conscious system because it lies in an unaware state most of the time, except when brought to awareness by a volitional act. Using this form of logic, Kubie arrived at the notion of the preconscious system. The preconscious system concept, and its speed of ideational flow, is further supported by the fact that we can converse in an intelligent manner without becoming aware of all the grammatical rules, syntax, related areas of knowledge, etc.

Kubie (1961) further hypothesized that many creative acts and discoveries are arrived at in this unaware way. The data to support this is drawn from accounts of creative people who explain they are normally unaware of putting together ideas in new ways (insight). The consequence of this process is a stream of ideas constantly flowing just below the threshold of consciousness. These ideas operate in a type of heuristic fashion in terms of metaphorical and analogical constructs, each being dominant only in response to a volitional act.

The conscious system samples the incessant flow of preconscious ideas for validating reality and communicating products. Metaphorically, it is a bridge connecting the private inner world with that of external reality. Creativity is thus defined as the interaction between
these two systems.

Still, another force operates to inhibit creativity, thus limiting the phenomena of creative thinking. This force is the unconscious, defined by Kubie (1961) as repressed thoughts. Due to the imperfection in repressing thoughts, certain thoughts "creep" into the preconscious system and disturb the balance between the preconscious system and the conscious system. This disturbance, or interference, causes a distortion of thought processes in the preconscious system and causes a neurotic process as a result.

Another example of a psychoanalytical approach to creativity is that of Fromm (1963). He distinguished two phases in dealing with productive or creative behavior. The first phase is essentially feminine in quality and refers to the conception phase of the product, analogous to birth giving. Following this phase, a more masculine act occurs during which the creator refines and polishes his work for social judgment. The relationship of this process is that in the creative act seemingly paradoxical entities are united, e.g., masculinity-femininity, producer-product, etc.

As previously mentioned, these two works (Kubie, 1961; Fromm, 1963) were presented to show how logical hypothetical constructs can be, yet even though elaborate, they fail to be empirically provable.

**Creativity as Personality**

The second area of research deals with creativity as personality. If a personality unique to the creative person could be discovered, more research questions could be asked and answered. MacKinnon (1962), working with creative persons at the University of California
at Berkeley, concluded that the following characteristics represent the group he studied. Creative people tend to function at a high level of effective intelligence. However, while high intelligence is indicative of general familiarity with existing knowledge, it is not necessarily indicative of being able to invent or innovate (Getzels and Jackson, 1962).

Another characteristic of the creative person was his openness to experience. Taylor (1963) pointed out that the cognitive style of creative people is characterized by humor, fantasy, and a playfulness with ideas. Radical ideas are entertained merely to see things differently, thus enlarging the realm of experience (Cattell and Drevdahl, 1955).

Other characteristics listed by MacKinnon (1962) included "freedom from crippling restraints and impoverishing inhibitions," aesthetic sensitivity, cognitive flexibility, and a high level of creative energy. Also mentioned was an intellectual style characterized by risk taking. This involves a type of readiness to accept the maximum amount of information one can receive from the external world (Kogan and Wallach, 1964). In so doing, the creative thinker often times will pay less attention to minute details but survey a broader spectrum of his environment. In attempting this, a higher risk of mistaken judgment is encountered, but a willingness to take a chance, to act impulsively, and to revise views previously held is characteristic of the creative person.

Other research seems to support the notion that creative thinkers are independent and non-conforming. Cattell and Drevdahl (1958) summarized creative thinkers as dominant, self-sufficient intellectually, radical and bohemian. While a precise definition of each of
these characteristics was not given, they obviously related to some form of independence and non-conformity, or unconventional behavior.

In summarizing the personality of the creative person, there was some agreement on the following stereotypic make-up. The personality of the creative individual usually consists of a high degree of effective intelligence. This effective intelligence is related to surveying the environment in a global manner, at the expense of observing minute details. In the process of excluding minute details, the creative person is more apt to be mistaken in his inferences. He takes risks, but mistaken inferences are only mistaken in the sense that they are different in terms of societal standards. Sometimes, these mistaken or reorganized perceptions lead to significant creative acts or products. Problems are looked at from a different perspective; an openness to new ways of looking at things is entertained (Gordon, 1961). The creative person is less concerned with what others will think about his actions and is then more free to synthesize sensory input in a different way. He has a type of sensitivity which allows him open or increased awareness to subtle cues in his general environment. The creative person is often times dominant in a group situation, yet usually prefers to be an introvert (Cattell and Dreydahl, 1958). These types of characteristics have been ascribed to creative persons. One must, however, keep in mind that these characteristics were presented in terms of group characteristics, and it is highly doubtful that any creative person would "fit the mold" in a precise manner. It should be noted that these conclusions are based on results from personality tests. A concern for the hypothetical nature, as well as vagueness of terminology arise. Another
source of this type of data are autobiographical and biographical in nature. Generally, one could question the bias of these types of reports.

**Creativity as a Product**

Creativity as a product, of necessity, deals with adjectives to define what type of product can be classed as creative. Such terms as originality, novelty, appropriateness, usefulness, etc. have been used to define this product. Jackson and Messick (1965) dealt with this area by defining what they considered to be the properties of a creative product. What was delineated was not proposed in terms of an absolute, but in terms of relativity to the setting in which the judgment of the creative product takes place.

They began by admitting that the most common and often times sole criteria of a creative product is its novelty or unusualness. Such questions were raised as to what compromises the set of objects by which the product in question is to be judged, such as subsets, universal sets, etc. A possible solution to this question is the establishment of norms. However, the authors maintained that, as a first step, the mere oddities be weeded out. Appropriateness was added as a second criteria to accomplish this process. "To be appropriate, a product must fit its context. It must make sense in light of the demands of the situation and the desires of the producer. Further, when products are complex, their internal elements must also blend together and be appropriate to each other, (Jackson and Messick, 1965, p. 313)." The role of this criterion was to eliminate from the set of unusual products the absurd, e.g., schizophrenic, psychotic responses which are often unusual, but many times inappropriate.
They also maintained that a further step in relating the concept of appropriateness to the product is to relate it in terms of the producer's intent.

According to Jackson and Messick (1965), another property necessary to judge a product as creative is the way in which its elements defy tradition and yield a new perspective, almost forcing the viewer to see reality in a new way. This power of transformation of reality is of a non-dichotomous nature (continuous rather than discreet), and Jackson and Messick wrote in terms of the powerfulness of the transformation. It must again be pointed out that it is quite difficult, if not impossible, to measure this power of transformation. Who is to determine the norms of reality and whether these norms have indeed been transformed?

These three criteria—unusualness, appropriateness, and transformation of reality were defined by the authors in terms of the viewer's response as surprise, satisfaction, and stimulation. These are all forms of esthetic responses in terms of the viewer.

Finally, Jackson and Messick listed a fourth quality of the creative product. This was condensation. Condensation refers to a type of non-acclimation to the novelty of the product. The greatest creative works of man have endured in a sense of timelessness. The authors contended that this is similar to the first impression of a creative product being simple, but when further examined, a type of complexity is evident. The reverse can also hold true, i.e., the first impression is complex but upon further inspection, a type of simplicity is perceived. These products "have about them an intensity and a concentration of meaning requiring continued contemplation, (Jackson and Messick, 1965, p. 320)." Like many, the above article was
hypothetical rather than empirical in nature.

Another inquiry into creativity as a product was that of Gordon (1961), in a book entitled Synectics. This approach was concerned with producing a creative outcome by bringing together a group of persons with diverse backgrounds, allowing for a diversity of responses to bear upon a singular problem area. Gordon's work is relevant to creativity as a product in the sense that all of the problems posed in this process are technological in nature, and the product can be termed creative or not by a sole criteria. The product works, or it does not work. This is a very limited yet quite pragmatic solution.

Research Studies Most Related

To Proposed Investigation

Few published research studies using an experimental analysis of behavior to examine divergent responding exist. The earliest study using this approach appeared in 1972, and since then, only three other studies were found. Since a dearth of research exists in a behavioral approach to divergent responding, a detailed presentation of each of these studies is presented.

Two studies were published in 1972, one by Ward, Kogan, and Pankove, and the other by Goetz and Salmonson. While both studies can be classed as behavioral, some exception must be made to the Ward, Kogan, and Pankove study.

As the title "Incentive Effects in Children's Creativity" suggested, Ward, et. al, examined environmental effects on what they called creativity. At this point, it is reasonable to assume this to be a

---

1. William Sharpe supported that genius was a function of environment as early as 1755, in his Dissertation Upon Genius.
behavioral model. Upon closer examination of the study, however, one finds that while behavioral, the authors chose to use a group design. This is immediately indicated by their tabular presentation of scores which reflected group means and standard deviation scores. As the study is presented, suggestions as to how changes in the research design used may have yielded differing interpretations of the data are given.

The problem as stated by Ward, et. al. (1972) revolved around the complexity and inconsistency of results obtained by various researchers in studying differences in creativity testing conditions; specifically, a game-like testing condition was compared with a more rigid testing situation. The authors cited examples of such studies by Boersma and O'Bryan (1968) and Kogan and Morgan (1969). From these inconsistencies they deduced that a simpler study investigating incentive variables would be appropriate. "The central concern was with the consistency of performance—-are relative individual differences in ideational production stable over variations in the motivational context of testing administration (Ward, et. al., 1972, p. 670)?"

One finds that the authors were concerned mainly with capacity of the individual, and environment or incentives acting on that individual. It seems as if they were proposing to answer a "nature vs. nurture" type argument in regards to what they call creativity. While the authors concluded that creativity is a function of capacity, it was felt by this writer that this conclusion was unwarranted in light of the findings proposed.

The subjects were fifth grade children, 95 males and 96 females, predominantly urban black children of lower socio-economic status. The authors reported the mean age as 133.3 months (SD=5.4 months).
Referring to IQ scores on the Kuhlman Anderson test, the authors reported that 66.66% of the subjects had a mean IQ score of 94.0 (SD=11.5). These scores were already recorded on the subjects, and the remaining 33.33% of the subjects were not given an IQ test. It is interesting to note that after running correlational tests on the subjects' creativity scores and IQ scores, the authors concluded that there is essentially no positive correlation between IQ scores and creativity scores. This conclusion, while supportive of other research finding the same results, e.g., Wallach and Kogan (1965), is unwarranted considering only 66.66% of the sample was used. It is feasible that the remaining third of the subjects could have altered the results. This is especially true considering that the IQ score and its relationship to creativity is inconclusive. It might be suggested that this problem could have been reduced by administering IQ tests to all of the subjects under similar conditions. This would have lengthened the time involved in the study, but this seems warranted since this issue was of major concern to the authors.

The tasks used in the Ward, et al., study (1972) were four measures consisting each of a four item modification of the original Wallach-Kogan Creativity Test (Wallach and Kogan, 1965). Two of these measures contained verbal content, one in which a child was asked to name uses for an object, and the other in which the child was asked in what ways a pair of objects were similar. The two remaining measures were figural in content in which the child was asked to interpret a simple abstract pattern and a line drawing. No time restrictions were placed on the responses.

Session one consisted of baseline information for all subjects, and the total group was then divided by sex and randomly assigned to
one of three groups. Two of the four measures were given under baseline conditions, one verbal and one figural. The remaining two measures were given during session two.

The results indicated a .12 average correlation, with a range of .04 to .17, for correlations between the number of ideas given by the subject and his IQ score. As previously mentioned, only 66.66% of the sample actually had IQ scores with which to make the correlation. Also, according to the authors, there were no sex differences. The results also compared favorably to that of middle class socio-economic children. Correlations between IQ scores and the quality of ideas, and quantity of ideas and quality of ideas, yielded no significant correlations statistically.

After converting the subjects' scores to a mean of 50, and a standard deviation of 10, and conducting an analysis of covariance, the authors reported that both experimental groups increased idea output when using immediate and delayed rewards as compared to the control group (p < 0.01).

A final correlation analysis was performed to determine if the ideational output over the various tasks remained stable. The authors reported the average correlation of the three groups as .60, .53, and .65, and concluded that no statistically significant variability existed across experimental conditions. The concluding statement of the authors was, "The alternative explanation—that the observed individual differences reflect variation in the capacity for divergent ideational production—thus receives support (Ward, et al., 1972, p. 675)."

This conclusion overlooks some basic logical, as well as methodological concerns. First, the authors wrote about individual differ-
ences and then relied on group scores and methods to measure these differences. If these differences do exist as a function of capacity, it would seem more practical to work with individuals and to compare each individual with himself rather than to compare a set of group scores with another. Then, if gains are noted, one could hypothesize functional relationships between the treatment and the gain scores. This could be accomplished using a reversal single organism design.

Another concern lies in the conclusion drawn from the study. Admittedly, money is a strong reinforcer in our society, but to assume, as the authors did, that one penny for each idea would be reinforcing is questionable empirically. No research has yet found a universal reinforcer, yet the authors proceeded as if this were true. Probably the most reasonable explanation of the absence of a universal reinforcer is due to individual differences. Therefore, to ask if individual differences exist as a function of capacity or incentives, and then to not allow for those differences in the design used may bias the results. In all fairness to the authors, pennies could be classed as interim reinforcers (Stephens, 1970) or tokens, and this would expand the individual's use of the reinforcers to purchase whatever he wished. However, using the authors' data, which stated that each child scored 17 to 20 ideas for each four item task, one finds that a child scoring as high as twenty ideas for the four measures could gain 80 cents for the total experiment. One might ask at this point in our history, just how reinforcing is 80 cents. To some children it may be more than sufficient, to others, not. This whole notion of what is reinforcing to the child could have been alleviated with a short assessment of reinforcers, or by using tokens in conjunction with a reinforcement list. With 191 subjects (the number used with Ward, et. al., 1972) assess-
ment would be quite cumbersome, and this could have been overcome by using a smaller number of subjects, five or ten. If, with a smaller number of subjects, it was determined that money was reinforcing, then more money could have been earned, possibly increasing idea output to higher levels.

Another advantage of a single organism design would have been the opportunity to vary the reinforcers until a change was noted. Certainly, if after a multitude of reinforcers were tried, and no change was noted, one could at least have greater evidence for supporting the "capacity" theory of divergent responding, if only for the particular subject under consideration. Then, one could seek replication of this phenomena with other subjects.

Goetz and Baer (1973) discarding the term creativity for the term divergent responding, argued that blockbuilding behavior was considered important as it is significantly a part of preschool programs. This behavior has not been empirically tested to determine its value in the schools, and the authors contended that if blockbuilding behavior did have value empirically, the value would lie not in tedious, repetitive, or stereotypic types of blockbuilding behavior, but in divergent blockbuilding behavior.

The subjects used in the Goetz and Baer study were three girls, each four years old, and having deficits in blockbuilding behavior. These deficits were specifically defined as either not building the blocks upon each other (e.g., just laid out in groups), or repetitive in nature (e.g., building the same castle over and over again).

In defining divergent behaviors, 20 general forms of blockbuilding behaviors (products) were arbitrarily identified, e.g., fence forms defined as "any two or more placed side by side in contiguity;...
(Goetz and Baer, 1973, p. 211)." A form diversity score was defined as the appearance of a defined form never before displayed by a given subject during the total observed sessions.

The procedures used by Goetz and Baer consisted first of baseline data collection. During this time (3 to 5 sessions) no criticism or enthusiastic statements were made to the subject. After the total session, the child was thanked for participating. At this time, the child was escorted outside and the teacher returned to photograph the block forms so they could be examined by independent judges and scored appropriately. Inter-observer reliability for the judges and the teachers was maintained at 95% or better throughout the study. After a stable baseline occurred, the subjects were reinforced with descriptive verbal praise such as, "Oh, that's very nice--that's different," for constructing a block form not previously displayed in that session. This phase was continued until clear evidence of increasing form diversity was obtained.

After reinforcing form diversity, a reversal of the verbal contingency marked the third phase of the study. This time, form sameness was reinforced verbally and descriptively. These sessions were continued until a decrease in form diversity was shown, approximating baseline level. The final phase, in order to leave the subjects at a higher level of diversity in blockbuilding behavior, was a reintroduction of the reinforcement for form diversity contingency.

The results indicated that the control of form diversity and sameness by verbal contingencies was established. Data was also presented to show that new forms appeared contiguously when reinforcement for form diversity was in effect.

One should note that previously this writer stated that ethical and
philosophical issues underlie various approaches to science. Several questions of this nature could be asked of the research just cited. First of all, the criteria of "diverse responses" was determined by the experimenter (analogous to society). Therefore, if diverse responses are predefined, they tend to oppose the concept of novelty, or originality of behavior, often times cited as an aspect of the concept of creativity. However, one could argue, as Goetz and Baer might, that novelty and originality can be said to exist only in terms of behavior displayed, as that is all that is open to quantification. While this approach may tend to negate the poetic but often times vague concept of creativity, it certainly allows the concept to be examined by the tools of science man now possesses, namely, quantification of data. In other words, the concept of creativity needs redefining, in terms of how science can investigate it. Of course, one extreme issue would be that ultimately the subjects would merely be reinforced for conforming to the experimenter's predefined goals. While this point will not be considered in depth in this paper, it is one to be remembered throughout this type of research. Goetz and Baer (1973) argued, however, that only descriptive reinforcement was given, such as sameness and diversity, and not instruction on how to construct a given form. One question that becomes immediately apparent is what if a subject builds a form not within the predefined forms of the experimenter? A possible answer to this question would be to add that particular new form to the list of predefined forms. To do this would raise other questions however; is it an appropriate form?, etc. So, a solution is far from being found.

Maloney and Hopkins (1973), also using divergent responses as one aspect of creativity, supported the earlier findings of Goetz and
Baer (1973) and Goetz and Salmonson (1972). That is, a variety of responses can be increased using contingencies established by the experimenter. The response class used in this study was sentence structure, and as well, its effect on subjective judgments of creativity in writing.

The subjects were fourteen elementary school children attending a volunteer remedial summer program. Prior to each daily session, a noun was randomly selected without replacement from a list of thirty nouns. This noun was written on the blackboard, and provided the topic for a ten sentence story. Upon completion, or after a forty minute time constraint, each subject turned in his story and returned to his assigned work. The experimenter then scored each paper and the number of points the child earned was written beside his or her name on the blackboard, ten minutes prior to recess.

The students were scored on a list of objective variables explicitly defined by the experimenter. An independent scorer was used, and he was not informed of the purposes of the experiment, or the design used. An example of a compositional variable was, "Number of words. Any group of letters, written or printed, representing a spoken word. Misspelled words were counted as words (Maloney and Hopkins, 1973, p. 426)."

Inter-observer reliability checks were made on the time to completion of a task, and each compositional variable was scored. The mean range on all reliability ratings for all measures was 78% to 97%.

During baseline, the noun was placed on the board, and each student handing in a paper with a ten sentence story received a score of 100. This score was placed on the blackboard.
After baseline, a game like condition was employed and consisted of two teams competing for points in order to receive recess five minutes earlier, and a piece of candy. The teams were determined by pairing the subjects on scores based upon baseline products. One point was given for each adjective, adverb, action verb; different adverb, adjective and action verb; different sentence beginning; compound sentence; prepositional phrase; and a sentence with more than eight words. It might be noted that the subjects were unaware of this scoring, for purposes of team assigning, and while they each ostensibly received 100 points for completing the assignment, each was actually grouped according to the above scoring procedure. After the subjects were paired, each was randomly assigned to a team.

Prior to writing the stories, each team was told each member would be given an arbitrary amount of points for using experimenter designated compositional variables. These variables shifted in emphasis after three sessions. Examples of the variables (e.g., twenty different adjectives were worth 100 points) were elicited from the class, reiterated by the experimenter, and written on the blackboard as an example. These examples were erased before the topic noun and statement of the contingency were announced to the teams.

Besides scoring papers, the independent judges were asked to rank the "creativity" of the papers. They were informed not to base their creativity rankings on spelling, punctuation, handwriting, or the topic of the story. Additionally, the raters were instructed only to rate the papers in comparison with other papers by the same subject, or on an intra-subject basis. The reliability of the ratings on creativity was 46%, or 25 agreements out of 54 stories. Chance re-
liability would be 14 out of 54 agreements, according to the authors, significant by use of a chi square test at p < .0001.

An examination of the mean results showed an increase in usage of compositional variables contiguous with the contingency. Also, while not as large an increase as when one contingency was in effect, the products, when all three contingencies were in effect (adjectives, action verbs, and sentence beginnings), increased each compositional variable over the baseline rate. In addition to mean results, individual analyses were presented, and eight subjects were very similar to the mean performance, and the remaining subjects had other increases over baseline levels besides the specific contingency named (generalization across conditions). Only one subject had mean scores on all variables remain constant throughout the study.

In analyzing the creativity rankings, the number ranked most creative under baseline conditions was one, while the rankings for least creative was 13. This relationship shifted as the contingencies were incorporated with the highest shift being during the different action verb contingency. During this contingency the most creative ranked were 16 in number, and the least creative were three in number.

The study just presented and that of Goetz and Baer (1973) had several commonalities. First, both studies, besides using an applied behavior analysis approach, explicitly used divergent responding as one aspect of the concept of creativity. Definitions of behaviors prior to the experiment were also determined. These definitions of observable behaviors are important if replications are to be made. Inter-observer reliability ratings are necessary to maintain the accuracy of the measurements.

A second characteristic of both studies was that the authors of
both were concerned with intra-subject controls. Goetz and Baer (1973) used comparisons based on such definitions as behaviors never before displayed by a given subject within the experimental setting, e.g., form diversity. Maloney and Hopkins (1973) similarly requested the independent raters to compare creativity in story writing only with the same subject's products. If individual differences do exist, there is a need to compare a subject with himself rather than to compare him to a group mean, to allow the variability of individual differences to be controlled. This line of argument was expressed by Sidman (1960).

Goetz and Salmonson (1972) investigated the relationship between diversity in easel painting and general and descriptive reinforcement. The subjects used in this study were three preschool children displaying a fairly low number of different forms when easel painting. (It should be noted that Goetz and Baer, 1973, also chose as subjects children who had deficits in a particular defined behavior. The notion of applied research enters into these types of studies as they are concerned with increasing deficit behavior in specific children.)

The subjects were asked to easel paint with the colors red, blue, and yellow. The subjects determined the length of the session by ceasing to paint.

Prior to the experiment, an analysis of paintings by preschool children was conducted and 25 painting forms were defined. As in the previous behavioral studies reviewed, the subjects could then act as their own controls, even though the painting forms were identified using group data.

The concept of descriptive reinforcement versus general reinforcement was dealt with as the contingency and was similar to that
of Goetz and Baer (1973). That is, social reinforcement like, "That's nice" was considered as general reinforcement, while reinforcement like "That's nice the way that line goes back and forth--like a zig-zag" was considered as descriptive reinforcement.

The authors stated that inter-observer reliability checks were made on several measures, with inter-observer reliability remaining at 99% or better throughout the study.

The conditions varied in sequence to control for the possibility of ordering effects. These conditions were baseline, general reinforcement, and descriptive reinforcement. The results were similar for all conditions, regardless of order, with descriptive reinforcement being slightly higher than general reinforcement, and baseline being less diverse than either of the treatment conditions.

Summary

The studies presented under creativity as process and creativity as personality mainly considered creativity as an independent variable. That is, creativity was an assumed concept, and the research attempted to define the characteristics of creativity, either as a process, or as a personality.

The studies presented under creativity as a product, especially those using a single organism design, dealt with creativity (or one aspect of creativity) as a dependent variable. Greater emphasis was placed on this approach as it is directly related to the present investigation.
CHAPTER III

METHOD

Subjects

Six subjects were used in the present study. Two subjects were selected (with principal recommendations) from an average population; two from a learning disabled population; and, two from a gifted population. The age range of the subjects was from 10-1 to 12-3 years, and all subjects were drawn from a public school setting. While the particular school used did not contain a special class for the learning disabled, the principal's recommendation based upon his knowledge of that student, as well as that student's school records, was used to classify these children. When an IQ score was not available in the student's school records, the Peabody Picture Vocabulary Test was administered by the experimenter.

Setting

The setting used in this study was a room in a public school not being used for classroom instruction. The subjects reported to this room at the end of the school day. To insure the attendance of the subjects, they were each paid a small stipend contingent upon completion of the study. The stipend was presented at the end of the study. The subjects were then seated across from the experimenter at a table void of objects, except for the recording instruments of the experimenter, i.e., pen, recording form, and test items.
Instrument

The Wallach-Kogan Creativity Test (1965) was designed to measure two variables of creativity, namely, uniqueness, and the total number of responses given to a single stimulus. The total number of responses was referred to as the numbers score. The test was designed to be given in a game-like atmosphere. This game-like structure was based on the premise that a structured testing situation may be antagonistic to the free flow of associations needed to respond in an appropriate manner, e.g., fluently. The test consisted of five subtests which were: 1) instances, 2) Alternate Uses, 3) similarities, 4) pattern meanings, and 5) line meanings.

Instances: In this subtest the child was asked to respond with numerous answers to questions involving class concepts, in verbal terms. An example was to respond to a statement like, "Name all the round things you can think of". The responses were counted to determine the numbers score.

Alternate Uses: In this subtest, also verbal in nature, the child was asked to name as many possible uses as he could for a particular object. A sample item was, "Tell me the different ways you could use a newspaper".

Similarities: This category referred to asking the child to say a number of ways in which two objects are alike. A sample statement in this subtest was, "Tell me the ways in which a potato and carrot are alike."

In scoring each subtest, the numbers score was calculated for each item of the subtest and then summed across items for that particular subtest. A complete list of items to each subtest are
given in Appendix A.

While Wallach and Kogan (1965) included the subtests of pattern meanings and line meanings in their test battery, they were not used in the present study. The reason for this omission was largely due to the treatment used in the study. The three subtests previously mentioned consist of a total of 22 items. This was a large amount of items for a subject to respond to in a given session, and since the intent of this investigation was to determine the effects of a reinforcement system upon divergent responding, the purposes of the study were not violated. For the interested reader, these two subtests are also presented in Appendix A.

As previously stated, the testing situation was as non-structured as possible while still having the subjects emit appropriate responses to the test items. Wallach and Kogan (1965) hypothesized that a structured testing situation was not conducive to responding appropriately in terms of divergent responses, i.e., fluency. The notion of restraints hindering divergent responding was hypothetical, and though this non-structured technique was originally used by Wallach and Kogan (1965), it was questioned as essential by Ward, Kogan, and Pankove (1972). They based this question on the premise that if reinforcers were given for responding, then the situation was no longer game-like, hence structured. While this may be semantically true, pragmatically it is irrelevant. This writer feels that if no time constraints were placed on the child's responses, a free and game-like atmosphere can prevail even if contingencies were placed upon the responses.

General instructions for administering the test items were given by Wallach and Kogan (1965), and can be examined in Appendix B.
It was also suggested by the authors that the items should be presented in a manner of suggestion rather than finality, indicating to the child that the answers are innumerable. Additionally, in the present investigation inter-observer reliability checks were made once during the baseline condition on each subject, and once during treatment for each subject. This was done to insure objectivity in judging a response as appropriate (fitting the context of the question).

**Design**

The design employed in this study was a combination of three behavioral designs. These were an intra-subject multiple baseline design, an inter-subject multiple baseline design, and a reversal design. The inter-subject multiple baseline design was used with two subjects from each of the previously defined categories, e.g., learning disabled. The reversal design was used with each subject individually, as also was the intra-subject multiple baseline design. This combination design is presented graphically in Figure 1. The use of all three of these designs was used to control for extraneous variables, as well as for replication of effects (Wolf and Risely, 1969).

**Baseline I**

The subjects were placed under two conditions during the present study, i.e., treatment and non-treatment. The first condition was Baseline I. During this condition the subjects were presented the three verbal subtests of the Wallach-Kogan Creativity Test. Each response was recorded on a recording form as shown in Appendix C. No reinforcement was given during baseline conditions.

After the session, the experimenter summed the responses for
Fig. 1. Graphic representation of design.
each subtest. A baseline session consisted of all the items in each of the three subtests presented to the subject as stimuli. The subject determined when a new item was to be presented by indicating that he had no additional responses to give to that item. If the experimenter noted a lag in responding he asked the subject in a matter of fact manner, "Is that all?".

The number of sessions in the Baseline I condition was determined by stabilization, or a descending slope in the data. Stabilization was considered as a series of scores showing little or no increase in numerical value. At this point baseline was terminated.

Treatment

Treatment was introduced to the first subject of each pair after stabilization, or a decreasing slope in the data occurred on the Instances subtest. Simultaneously, the second subject of each pair was continued on baseline for all three subtests. Treatment was introduced on only one subtest at a time. Treatment remained in effect until a continued increase over baseline levels occurred. The purpose of treating only one subtest at a time was to adhere to the intra-subject multiple baseline design. After a continued increase was noted on the Instances subtest, treatment was removed from that subtest. Treatment was then introduced on the second subtest, or the Alternate Uses subtest for the first subject of the pair. At the same time, treatment was introduced for the first subtest, or the Instances subtest for the second subject of the pair. As each of these treated areas showed a continued increase, treatment was removed from that particular subtest and introduced on the next subtest. After all of the subtests were treated for the first subject
treatment was reintroduced for all three subtests simultaneously for one session.

The treatment consisted of a contingency allowing the subjects to "earn" interim reinforcers which could then be traded for tangible rewards or high interest activities. To increase the probability that the tokens would function as a reinforcer, the activities and rewards were listed in a menu type format containing the prices of each of the items.

A session during the treatment condition consisted of all items of a particular subtest being reinforced, as well as the other two subtests which were not being reinforced. The subject determined when he was finished responding to an item.

**Tokens and Exchange Period**

The tokens used in this study were plastic poker chips. These tokens were dispensed immediately following a response, and were placed next to the subject. The tokens were dispensed on a continuous reinforcement schedule at a rate of one per each appropriate response. The tokens were also paired with verbal praise such as, "That's a good answer."

The subjects were asked to allow the tokens to accumulate, and at the end of the session the child was shown a list of reinforcers with their corresponding value. The child was also told that if he wished he could allow the tokens to accumulate across treatment sessions and the number would be recorded in a "bank book". The tokens could then be spent at a later date. The "bank book" consisted of a folded piece of paper which showed the total tokens earned, those
spent, and the balance remaining. An example of the "bank book" is shown in Appendix D. The experimenter kept a duplicate of the "bank book" to assure the accuracy of the books. When the child decided to spend the tokens he was shown the reinforcement list. This list is shown in Appendix E.

Reintroduction of Reinforcement Contingencies

After all the subtests were treated and reversed for the first subject of each pair the reinforcement contingencies were reintroduced for all three subtests simultaneously for one session. This was done as an intra-subject replication to support the functional relationship between the independent variable (token reinforcement) and the dependent variable (frequency of responses given to stimulus items).
CHAPTER IV

RESULTS

The results presented in graphs I to III indicate an increase of responses to stimulus items when treatment was introduced. That is, the frequency of responding was higher for all subjects when the reinforcement contingency was in effect. These increases were noted for all of the subtests given. When all three subtests were treated simultaneously, a subsequent increase was noted for all three subtests. Inter-observer reliability was 99.3% for appropriateness of the responses, throughout the study.

Gifted Group

Subject G

The first subject of this pair had a mean baseline score of 29.66 for the Instances subtest (three sessions), 30.71 for the Alternate Uses subtest (seven sessions), and 49.00 for the Similarities subtest (10 sessions). During the treatment condition, the Instances subtest mean scores increased to 187.50, the Alternate Uses score to 120.00, and the Similarities subtest score to a mean of 110.30.

The number of sessions for each treatment condition were four sessions for treatment on the Instances scores, and three sessions for treatment on the Alternate Uses and Similarities subtest scores.

During the reversal condition, the mean scores for each subtest decreased to 45.17 for the Instances subtest, and 37.00 for the Alternate Uses subtest. No reversal was obtained for the Similarities subtest.
When treatment was introduced simultaneously for all three of the subtests, all three scores increased again. The score of the Instances subtest increased to 201.00; the Alternate Uses to 120.00; and the Similarities to 150.00. These scores were obtained during one session, with the reinforcement contingency in effect for all three subtests. (The mean scores for each subject on each subtest are summarized in Table I).

Subject $G_2$

During baseline conditions the mean scores for this subject were 23.71 for the Instances subtest (seven sessions), 32.82 for the Alternate Uses subtest (11 sessions), and 36.43 for the Similarities subtest (14 sessions).

During the treatment condition, an increase to a mean of 183.00 on the Instances subtest was noted. The Alternate Uses subtest increased to 239.33, and the Similarities to 266.00. The number of sessions were four for the Instances subtest, and three for both the Alternate Uses and Similarities subtest.

Mean scores for the reversal conditions were 14.66 for the Instances subtest, and 24.66 for the Alternate Uses subtest. No reversal was obtained for the Similarities subtest. (See Table I. Graph I represents the scores obtained for all conditions for subjects $G_1$ and $G_2$).

Average Group

Subject $A_1$

This subject had mean baseline scores for the Instances subtest, the Alternate Uses subtest, and the Similarities subtest of 32.33 (three sessions), 28.00 (seven sessions), and 35.10 (10 sessions),
<table>
<thead>
<tr>
<th>Instances</th>
<th>Alternate Uses</th>
<th>Similarities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>$\bar{X}$ Baseline</td>
<td>$\bar{X}$ Baseline</td>
</tr>
<tr>
<td>$G_1$</td>
<td>29.66</td>
<td>30.71</td>
</tr>
<tr>
<td>$G_2$</td>
<td>23.71</td>
<td>32.82</td>
</tr>
<tr>
<td>$A_1$</td>
<td>32.33</td>
<td>28.00</td>
</tr>
<tr>
<td>$A_2$</td>
<td>49.29</td>
<td>37.33</td>
</tr>
<tr>
<td>$LD_1$</td>
<td>28.00</td>
<td>45.27</td>
</tr>
<tr>
<td>$LD_2$</td>
<td>42.33</td>
<td>31.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>$\bar{X}$ Treatment</th>
<th>$\bar{X}$ Treatment</th>
<th>$\bar{X}$ Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_1$</td>
<td>187.50</td>
<td>120.00</td>
<td>110.30</td>
</tr>
<tr>
<td>$G_2$</td>
<td>183.00</td>
<td>239.33</td>
<td>266.00</td>
</tr>
<tr>
<td>$A_1$</td>
<td>160.00</td>
<td>96.33</td>
<td>98.66</td>
</tr>
<tr>
<td>$A_2$</td>
<td>126.40</td>
<td>85.33</td>
<td>112.66</td>
</tr>
<tr>
<td>$LD_1$</td>
<td>212.25</td>
<td>224.33</td>
<td>230.33</td>
</tr>
<tr>
<td>$LD_2$</td>
<td>126.75</td>
<td>106.00</td>
<td>98.66</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>$\bar{X}$ Reversal</th>
<th>$\bar{X}$ Reversal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_1$</td>
<td>45.17</td>
<td>37.00</td>
</tr>
<tr>
<td>$G_2$</td>
<td>14.66</td>
<td>24.66</td>
</tr>
<tr>
<td>$A_1$</td>
<td>34.33</td>
<td>31.33</td>
</tr>
<tr>
<td>$A_2$</td>
<td>30.00</td>
<td>36.33</td>
</tr>
<tr>
<td>$LD_1$</td>
<td>14.00</td>
<td>25.00</td>
</tr>
<tr>
<td>$LD_2$</td>
<td>21.83</td>
<td>32.66</td>
</tr>
</tbody>
</table>

**TABLE I**

*Mean scores during experimental conditions.*
Graph I represents the frequency of verbal responses given to stimulus items on the Wallach-Kogan Creativity Test under the varying experimental conditions. Scores for all three subtests are shown.
respectively.

An increase in the mean scores for the three subtests during the treatment condition to 160.00 for the Instances subtest (four sessions), 96.33 for the Alternate Uses subtest (three sessions), and 98.66 for the Similarities subtest (three sessions), was noted.

For the reversal condition on the Instances subtest, the score dropped to a mean of 34.33; for the Alternate Uses subtest to a mean of 31.33. These scores approximated baseline conditions. No reversal was obtained for the Similarities subtest. (See Table I).

When treatment was introduced simultaneously for all three subtests, the scores increased to 180.00 for the Instances subtest, 81.00 for the Alternate Uses subtest, and 100.00 for the Similarities subtest (one session). (Graph II represents the scores for all conditions for both subjects $A_1$ and $A_2$).

Subject $A_2$

Subject $A_2$ obtained a mean score for baseline conditions of 49.29 for the Instances subtest (seven sessions), 37.33 for the Alternate Uses subtest (12 sessions), and 48.07 for the Similarities subtest (15 sessions).

During treatment conditions an increase to a mean of 126.40 for the Instances subtest was noted. The Alternate Uses subtest score rose to a mean of 85.33, while the Similarities subtest increased to a mean of 112.66. Five sessions were allotted to treatment on the Instances subtest, and three sessions each to the Alternate Uses and Similarities subtest.

Reversal conditions were related to a decrease of mean scores on all subtests. The Instances subtest dropped to 30.00, the Alternate Uses subtest to 36.33, and no reversal was obtained for the Similari-
Instances = o—o
Alt. Uses = x—x
Similarities = * *

Subjects $A_1$ and $A_2$

Graph II represents the frequency of verbal responses given to stimulus items on the Wallach-Kogan Creativity Test under the varying experimental conditions. Scores for all three subtests are shown.
ties subtest. (See Table I and Graph II).

**Learning Disabilities Group**

Subject LD$_1$

This subject obtained a mean score during baseline conditions of 42.33 for the Instances subtest (three sessions), 31.00 for the Alternate Uses subtest (seven sessions), and 29.80 for the Similarities subtest (10 sessions). (See Table I).

During the treatment condition an increase to a mean of 126.75 for the Instances subtest (four sessions) was noted. The Alternate Uses subtest increased to a mean of 106.00 (three sessions), and the Similarities subtest increased to a mean of 98.66 (three sessions).

Upon reversal, a drop was detected in the mean scores on all of the subtests. The Instances subtest dropped to a mean of 21.83, and the Alternate Uses subtest dropped to 32.66. No reversal was obtained for the Similarities subtest.

When treatment was reintroduced for all three subtests simultaneously, increases to scores of 185.00 for the Instances subtest, 123.00 for the Alternate Uses subtest, and 157 for the Similarities subtest were noted (one session). (Graph III represents the scores for all conditions for subjects LD$_1$ and LD$_2$).

Subject LD$_2$

This subject obtained mean scores for baseline conditions of 28.00 for the Instances subtest (seven sessions), 45.27 for the Alternate Uses subtest (11 sessions), and 48.93 for the Similarities subtest (14 sessions).

During the treatment condition these scores increased to a mean of 212.25 for the Instances subtest (four sessions), 224.33 for the
Graph III represents the frequency of verbal responses given to stimulus items on the Wallach-Kogan Creativity Test under the varying experimental conditions. Scores for all three subtests are shown.

Instances = o—o
Alt. Uses = x—x
Similarities = *—*  Subjects LD1 and LD2
Alternate Uses subtest (three sessions), and 230.33 for the Similarities subtest (three sessions).

During the reversal phase, the mean scores dropped to 14.00 for the Instances subtest, 25.00 for the Alternate Uses subtest, and no reversal was implemented during the Similarities subtest.

Statistical Analysis

A Kruskal-Wallis One Way Analysis of Variance by ranks was performed to test the null hypothesis. The null hypothesis was that there would be no statistically significant differences between the mean gain scores, across the varying categories of subjects, i.e., gifted children, learning disabled children, and average children.

No statistically significant differences were found between the mean gain scores of any of the groups (pairs of subjects). (See Table II).

Table II
Results of Kruskal-Wallis One Way Analysis of Variance by Ranks

<table>
<thead>
<tr>
<th>Subtest</th>
<th>H Value</th>
<th>P. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instances</td>
<td>1.99885</td>
<td>N.S. *</td>
</tr>
<tr>
<td>Alternate Uses</td>
<td>3.71305</td>
<td>N.S. *</td>
</tr>
<tr>
<td>Similarities</td>
<td>1.14175</td>
<td>N.S. *</td>
</tr>
</tbody>
</table>

* An H Value of 4.5714 was needed for statistical significance at the .067 level (Siegel, 1956, p. 282).
Summary

When treatment was introduced, the frequency of the number of verbal responses given to each stimulus item increased when compared to that subject's performance during baseline conditions. This increase in responding was noted for all subjects for all three of the subtests. Specifically, token reinforcement, when introduced contingent on diversity of verbal responses, increased the diversity of verbal responses (number of responses given to each item).

Additionally, no statistically significant differences were found in mean gain scores between each of the three pairs of subjects used.
CHAPTER V

DISCUSSION, LIMITATIONS, IMPLICATIONS AND SUMMARY

Discussion

Research Questions

The primary purpose of this study was to determine what effects token reinforcement would have on one aspect of the Wallach-Kogan Creativity Test. This variable was fluency, or the number of verbal responses given by the subject to a single verbal stimulus. Guilford (1950) and Wallach and Kogan (1965) refer to this type of divergent responding as one aspect of creativity. Goetz and Baer (1973), Maloney and Hopkins (1973), and Goetz and Salmonson (1972) used reinforcement to increase divergent responding in easel painting, blockbuilding, and story writing behaviors. The results of the present study (presented in Graphs I, II, and III) supported the findings of Goetz and Baer, Maloney and Hopkins, and Goetz and Salmonson. That is, token reinforcement has served to increase divergent responding in a given area, specifically, verbal fluency.

The secondary concern of the study, which was to determine if any statistically significant differences in mean gain scores existed between the varying categories of subjects, failed to show such. That is, the null hypothesis that no statistically significant differences in mean gain scores between groups would occur, must be accepted. It would be logical to assume that this non-significance could indicate that token reinforcement worked equally well with all of the groups used, relative to each subject's baseline level of performance.
Discussion of each category of subjects.

Gifted Group

The mean baseline scores of these two subjects varied approximately six points for the Instances subtest, two points for the Alternate Uses subtest, and 13 points on the Similarities subtest. The discrepancy noted on the Similarities subtest could be due to an extinction effect on the part of the second subject of that group (subject G₂). This effect could be a function of extinction due to the extended baseline condition for subject G₂, which was 14 sessions as compared to 10 sessions for subject G₁. However, subject G₁ obtained several extremely high scores compared to subject G₂ during baseline conditions. The highest Baseline I score for subject G₂ was 53, while that of G₁ was 63.

Additionally, a type of stimulus generalization could have occurred for the treated subject. That is, as one subtest was treated, a slight increase occasionally occurred on the following subtests which received no reinforcement. While this was not consistently shown as an overall increase in scores of the untreated subtests, the gains are noted only at the initiation of treatment on the first subtest. This could easily be considered as a stimulus discrimination paradigm (Reynolds, 1968). That is, initially, stimulus generalization occurred, as verbal responses on the previous subtest were reinforced. However, as the sessions progressed, differential reinforcement was affecting the rate of the responses of the subjects to the particular verbal stimulus item.

Mean scores under treatment conditions varied three points on the Instances subtest, 119 points on the Alternate Uses subtest, and 156 points on the Similarities subtest. The large difference on the last
two subtests would suggest a great deal of variability in the power of the reinforcers on each subject.

Reversal curves for both subjects shown on Graph I (p. 47) are representative of extinction curves following the withdrawal of continuous reinforcement (Ferster and Perrot, 1968).

Average Group

Greater variability in terms of inter-subject comparisons existed in this group. The mean baseline scores for subjects A₁ and A₂ differed by 17, 9, and 13 points on the Instances, Alternate Uses and Similarities subtests respectively. The differences during treatment conditions were 34 points on the Instances subtest, 11 points on the Alternate Uses subtest, and 14 points on the Similarities subtest. Reversal conditions were similar to those of the gifted group (See Graphs I and II, pp. 47 and 49).

Learning Disabled Group

Referring to Graphs I, II, and III, one can note much the same general pattern of response increase during treatment as compared to baseline. Greater increases can be noted with subject LD₂, but these greater increase values were not statistically significant when combined with subject LD₁. Subject LD₁ displayed an interesting example of how extraneous variables can affect behavior. During her first session of treatment on the Instances subtest she obtained a score of 41. Following this session she continually increased her verbal responses during each subsequent session of treatment. The contiguity of obtaining a score of 41 on her first treatment session and being sent to the principal after school could account for this. She was threatened with corporal punishment and came into the experimental session with tears in her eyes. After talking with her
for a while she agreed to complete the experimental session. Of
course, no empirical evidence is available to support any relationship
between her emotional state and her low score, but the incident does
lead to empirical questions concerning the isolation of some variables
but not being able to remove chance variables. Hence the need of
strong reinforcement is indicated for the subject to perform at only
baseline conditions.

Discussion of Statistical Results

The null hypothesis had to be accepted due to lack of statistical
significance as shown by the Kruskal-Wallis One Way Analysis of
Variance by Ranks. As previously stated, non-significance could be
construed as the token reinforcement working equally well with all
three groups of subjects (learning disabled children, gifted children,
average children). These results would substantiate the contention of
Wallach and Kogan (1965) that scores on their creativity test were
relatively independent of IQ scores. The range of IQ scores for the
subjects was 80 to 149 on either the Stanford-Binet Scale or the
Peabody Picture Vocabulary Test. It should be noted that to test
whether IQ scores and scores on the Wallach-Kogan Creativity Test
were relatively independent of each other was not the purpose of the
present investigation and therefore caution should be taken with the
above statement as no rigid controls were used, e.g., one cannot
simply substitute an IQ score on the Peabody Picture Vocabulary Test
for one on the Stanford-Binet Test with any valid comparison.

Limitations

While some limitations of the present study have been alluded to
throughout the above discussion, a need exists to summarize them in a coherent fashion, as well as discussing other limitations of this investigation.

Of some importance would have been to include other categories of subjects such as an educable mentally retarded population and possibly a trainable level retarded population. The advantage of this would be that if one were interested in the relationship between IQ score and Wallach-Kogan Creativity scores, the populations studied would span the continuum of IQ scores.

Another limitation to the study, though inherent in single organism designs, is that of immediate generalizations of the results. One must consider the size of the sample used and the possibility of those subjects representing extreme cases of each category. While one could conclude from the results that for the subjects used there did occur a positive relationship between an increase in numbers scores on all of the subtests and the introduction of the reinforcement system, to conclude that gifted children, average children and learning disabled children, in general, would be affected similarly by the use of a token reinforcement system would be unwarranted. For credible generalization to occur one would include the need for replication with other subjects from similar categories as those studied.

A third limitation of the study would be the effects of paying subjects for participating in the experiment. While this is commonly practiced in obtaining subjects for many psychology experiments in the university setting, and while this variable would not change the overall conclusions in terms of the effect of the reinforcement system, it may alter the stability level of the baseline scores. In other words, an empirical question could be asked concerning the effects of paying
subjects for participating in experiments and the resultant baseline levels. This could be controlled for by using another group of subjects not being paid for participation.

A fourth and final concern presented as a limitation to the present study is reflected in the categorization of the subjects. To label a child explicitly as learning disabled or educable mentally retarded, etc. is tenuous. Many times IQ scores, and principal and teacher recommendations are in error. Often this writer has questioned the certainty of the classification of the subjects used. While many researchers would consider this classification as necessary, some would contend that in an applied setting one is solely interested in changing the behavior of the subjects being worked with (Risely, 1969).

Implications and Recommendations for Further Research

Implications

Some of the implications which can be drawn from the present investigation have to do with a manner in which the concept of creativity can be researched. This would involve an approach to investigate relationships between variables of creativity, such as the variable used in this study, divergent responses. It does not, however, claim to bring consensus upon just what the concept of creativity is, or what its variables can be called. It does propose a process of defining variables of creativity into behaviors which can be observed, and then measuring the scientific manipulation of those variables. At present, no method is proposed to measure the covert behaviors which occur, and are associated with the concept of creativity. It may be useful to consider creativity on an individual level (Goetz and Baer, 1973). That is, creativity could be viewed as a lack of constructive diversity
of response patterns. This definition would not pacify all those who
do research in creativity, but it does not exclude a behavioral approach
as a viable approach to the topic of creativity. However, it must be
re-emphasized that this type of definition need not be all inclusive, as
surely it is not, of what has typically been defined as the concept of
creativity.

Another implication based upon the findings of the present study
would be that if verbal fluency, or divergent verbal responding, are
considered an aspect of creativity, then this aspect of creativity exists
with diverse types of children, at least those represented by the sub-
jects used. If this were true, or supported by further research, then
possible classroom implications could be sought. Children could be
assessed on a needs basis and a program implemented to increase
diversity in responding to environmental cues in a constructive manner.
One example of constructive diversity in responding would be for the
child to question the teacher when appropriate to facilitate that child's
learning responses.

Recommendations for Further Research

A series of studies can be proposed from the results of the present
study, including replication for generality of treatment. As was shown
in the graphic presentation of the results, one can note a marked
difference in the power of the token reinforcement on the individual
subjects by the amount of gain in frequency of verbal responding. This
observation could lead to a series of studies involving the relative
power of reinforcers among and within individual subjects. The power
of the reinforcer could be determined by the resultant effects upon the
behavior being measured.

Another investigation, related to the power of the reinforcer, is
the power of particular schedules of reinforcement upon divergent verbal responses to verbal stimulus items. Only a continuous schedule of reinforcement was used in the present investigation and one could certainly empirically question if other schedules would have had similar, greater, or lesser magnitudes of change than did the continuous schedule.

Still another investigation might involve the effects of reinforcement on other variables of creativity. Novelty, or the original occurrences of a particular behavior might be increased with reinforcement contingencies. This behavioral design could be extended to any observable behaviors which can be specifically defined and measured, and considered one aspect of creativity.

One might also investigate the effects of paying subjects for participating in experiments and its relationship to baseline stability and level. This was discussed previously in the limitations section.

A final recommendation would be to extend the present study to other populations such as the mentally retarded, etc. This could function as a means of further investigating the relationship between IQ scores and particular creativity measures, or variables. It would also begin to answer the empirical question of generality of effects to the population under study.

Out of the above recommended studies would, of course, grow more recommendations for further studies. In this manner, more empirical data can be obtained on creativity, and possible positive directions in creativity research could be established.

**Summary**

The present study examined the effects of a token reinforcement
system upon divergent responding to verbal stimulus items from the Wallach-Kogan Creativity Test. The primary purpose of the study was to determine if token reinforcement could increase what might be considered one aspect of the concept of creativity. That aspect was the number of appropriate verbal responses emitted to a single verbal stimulus item from a particular creativity test. A secondary purpose was to determine that if positive effects were observed during intervention, were the mean gain effects different across the categories of subjects used (gifted children, average children, and learning disabled children).

The dependent variable used in the present investigation was the number of verbal responses emitted by the subject. These were responses emitted to the verbal items from the Wallach-Kogan Creativity Test. Whether a verbal response was accepted as appropriate or not was determined by the experimenter. This was based on whether the response fit the context of the verbal stimulus. Inter-observer reliability was maintained at 99.3% throughout the study. The independent variable was the introduction of a contingency involving earning tokens for verbal responses (one token for each appropriate verbal response emitted). The tokens earned could then be traded for tangible items in the form of games, books, toys, etc.

The findings showed that when the contingency was in effect for earning tokens for appropriate responses, an increase over baseline levels was observed. Upon reversing the conditions, the response frequency decreased to levels similar to baseline conditions. This data increases the probability of a functional relationship between the introduction of token reinforcement and the number of verbal responses emitted by the subjects.
The findings related to the secondary purpose of determining statistical significance between any differences in mean gain scores between the categories of subjects revealed non-significance. In other words, this statistical non-significance of differences in mean gain scores between groups could be interpreted to mean that the token reinforcement used worked equally well with all of the pairs of subjects. This may be related to the flexibility applied setting research designs must have in order to account for the variability of human behavior.

It could be concluded that: (a) Token reinforcement contingent upon verbal responses to a given verbal stimulus increased the number of divergent responses given by the subject to that stimulus. This divergency in responses is considered to be one aspect of the concept of creativity. Caution should be used in terms of generalizing the findings to the populations represented due to the design used. (b) The results also supported the contention of Wallach and Kogan (1965) that IQ scores and creativity scores are relatively independent. This conclusion was drawn from the results, after the experiment was completed. Therefore, no rigorous controls were designed to test that particular relationship.
REFERENCES
REFERENCES


APPENDIXES
APPENDIX A

Items on the Wallach-Kogan Creativity Test

Instances

1. "Name all the round things you can think of."

2. "Name all the things you can think of that will make a noise."

3. "Name all the square things you can think of."

4. "Name all the things you can think of that move on wheels."

Alternate Uses

1. "Tell me all the different ways you could use a newspaper."

2. "Tell me all the different ways you could use a knife."

3. "Tell me all the ways you could use an automobile tire--either the tube or the outer part."

4. "Tell me all the different ways you could use a cork."

5. "Tell me all the different ways you could use a shoe."

6. "Tell me all the different ways you could use a button--the kind that is used on clothing."

7. "Tell me all the different ways you could use a key--the kind that is used in doors."

8. "Tell me all the different ways you could use a chair."

Similarities

1. "Tell me all the ways in which a potato and a carrot are alike."
2. "Tell me all the ways in which a cat and a mouse are alike."
3. "Tell me all the ways in which a train and a tractor are alike."
4. "Tell me all the ways in which milk and meat are alike."
5. "Tell me all the ways in which a grocery store and a restaurant are alike."
6. "Tell me all the ways in which a violin and a piano are alike."
7. "Tell me all the ways in which a radio and a telephone are alike."
8. "Tell me all the ways in which a watch and a typewriter are alike."
9. "Tell me all the ways in which a curtain and a rug are alike."
10. "Tell me all the ways in which a desk and a table are alike."

Pattern Meanings
Stimulus materials for the Pattern Meanings Test
Line Meanings

1

2

3

4

5

6
Stimulus materials for the Line Meanings Test.
APPENDIX B

General Instructions for the Wallach-Kogan

Creativity Test

Instances

"In this game I am going to tell you something and it will be your job to name as many things as you can think of that are like what I tell you. For example, I might say 'things that hurt.' Now you name all the things you can think of that hurt." (The experimenter then lets the child try.) "Yes, those are fine. Some other kinds of things might be falling down, slapping, bruises, or knife." (Here the experimenter varies her suggestions so that they consist of ones which the child has not provided.) "So we see that there are all kinds of different answers in this game. Do you see how we play?" (If the child already indicates strong understanding, the last sentence is replaced by, "I can see that you already know how to play this game.") "Now remember, I will name something and you are supposed to name as many things as you can think of that are like what I've said. OK, let's go."

Alternate Uses

"Now, in this game, I am going to name an object--any kind of object, like a light bulb or the floor--and it will be your job to tell me lots of different ways that the object could be used. Any object can be used in a lot of different ways. For example, think about string. What are some of the ways you can think of that you might use string?" (The experimenter lets the child try.) "Yes, those are fine. I was thinking that you could also use string to attach a fish hook, to jump rope, to sew with, to hang clothes on, and to pull shades." (The experimenter varies her suggestions so as not to duplicate any the child has provided.) "There are lots more too, and
yours were very good examples. I can see that you already understand how we play this game. So let's begin now. And remember, think of all the different ways you could use the object that I name. Here we go."

Similarities

"In this game I am going to name two objects, and I will want you to think of all the ways that these two objects are alike. I might name any two objects--like door and chair. But whatever I say, it will be your job to think of all the ways that the two objects are alike. For example, tell me all the ways that an apple and an orange are alike. (The child then responds,) "That's very good. You've already said a lot of the things I was thinking of. I guess you could also say that they are both round, and they have seeds, they both are sweet, they both are fruits, they both have skins, they both grow on trees--things like that. Yours were fine, too." (The experimenter's suggestions are varied so as not to include any which the child has given,) "Do you see how we play the game?" (If the child indicates clear understanding already, the last sentence is replaced by, "I can see that you already know how to play this game.") "Well, let's begin now. And remember, each time I name two objects you name as many ways as you can that these two objects are alike."

Pattern Meanings

"Here's a game where you can really feel free to use your imagination. In this game I am going to show you some drawings. After looking at each one, I want you to tell me all the things you think each complete drawing could be. Here is an example--you can turn it any way you'd like to." (The experimenter gives the example card to the child. See Appendix A,) "What could this be?" (The child is encouraged to try some suggestions). "Yes, those are fine. Some
other kinds of things I was thinking of were the rising sun, a porcupine, eye lashes, a brush, a carnation, and probably there are lots of other things too. And yours were very good examples too." (The experimenter's particular suggestions are varied so as not to include any given by the child.) "I can see that you already know how we play this game. So let's begin now."

Line Meanings

"This game is called the line game. I am going to show you some lines and after you have looked at each one, I want you to tell me all the things it makes you think of. Now take your time, and be sure that when you look at the line you tell me what the whole line makes you think of, and not just a part of it. O.K.?"

The experimenter then presents the first of the nine items in this procedure. Each line is shown on a separate 4 in. x 6 in. card. See Appendix A for a presentation of these stimulus materials. The experimenter now proceeds:

"Here is the first line. You can turn it any way you want to. Tell me all the things you can about it. What does it make you think of?"
APPENDIX C

SUBJECT__________ SESSION ________ DATE ________

SUBTEST: INSTANCES

RESPONSES:

1. 

2. 

3. 

4. 

SUBTEST: ALTERNATE USES

1. 

2. 

3. 

4. 

5. 

6. 

7. 

8. 

A sample of the recording form that was used in this investigation.
APPENDIX D

Bank Book

<table>
<thead>
<tr>
<th>Tokens Earned</th>
<th>Tokens Spent</th>
<th>Date</th>
<th>Balance</th>
</tr>
</thead>
</table>

Sample of subjects' "Bank Book" for saving tokens.
APPENDIX E

REINFORCEMENT LIST

<table>
<thead>
<tr>
<th>Item</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinner at MacDonald's</td>
<td>600</td>
</tr>
<tr>
<td>Baseball</td>
<td>500</td>
</tr>
<tr>
<td>45 RPM Record</td>
<td>300</td>
</tr>
<tr>
<td>Felt Tip Pen</td>
<td>200</td>
</tr>
<tr>
<td>Play-Doh</td>
<td>200</td>
</tr>
<tr>
<td>Comic Books</td>
<td>200</td>
</tr>
<tr>
<td>Ink Pen</td>
<td>150</td>
</tr>
<tr>
<td>Colored Pencils</td>
<td>150</td>
</tr>
<tr>
<td>Note Pad</td>
<td>100</td>
</tr>
<tr>
<td>Jacks</td>
<td>100</td>
</tr>
<tr>
<td>Checkers</td>
<td>100</td>
</tr>
<tr>
<td>Pencil</td>
<td>75</td>
</tr>
<tr>
<td>Other Choices (cost to be determined)</td>
<td>?</td>
</tr>
</tbody>
</table>