1979

The Bender-Gestalt test as a measure of creative production

Craig Logan House
Portland State University

Recommended Citation

10.15760/etd.2774

Title: The Bender-Gestalt Test as a Measure of Creative Production.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

Max R. Reed, Chairman
Walter G. Klopfer
Barry F. Anderson

In recent years, scientists have begun to turn their attention to research in areas long thought unsuitable to the scientific method. Research in the area of creativity has been notably absent until recently. Even though enjoying an increase in popularity, most of the research in creativity has produced contradictory results, and the methods employed to gather this data, in light of the interpretations, questionable.

This study attempted to define creativity in terms of objective, visible products of behavior. An existing psychological instrument, the Bender-Gestalt Test (BGT)
was modified in presentation and scoring methods into a test of creative production. This test was chosen because of its similarity to some existing tests of creative production and its theoretical relationship to the associative model of creative production. It was combined with the Making Objects Test (MO) and the Remote Associations Test (RAT) into a three test battery and administered to 90 college students. The three tests were scored and the results correlated with each other to determine the degree of relationship.

Significant correlations were found between the revised form of the BGT and the RAT, the RAT and the MO, and one judge's holistic rating and the RAT and the revised BGT. The results indicate that the revised BGT shows some promise, with future refinement, of becoming a quick, easy, and accurate method of psychological assessment of creative production potential.
THE BENDER-GESTALT TEST AS A
MEASURE OF CREATIVE PRODUCTION

by
CRAIG LOGAN HOUSE

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
PSYCHOLOGY

Portland State University
1979
TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of
Craig Logan House presented May 22, 1979.

Max R. Reed, Chairman

Walter G. Klopfer

Barry F. Anderson

APPROVED:

Robert E. Jones, Head, Department of Psychology

Stanley E. Rauch, Dean of Graduate Studies and Research
ACKNOWLEDGEMENT

I wish to express gratitude to all the members of the thesis committee for their concern, support and advice, particularly Dr. Max Reed, Chairman, and other members of the faculty of the Psychology Department who aided me in the development of this paper, particularly Dr. James Paulson. In addition, I would like to dedicate this paper to my wife, Sharon, whose unending support and encouragement helped me complete this work despite many obstacles.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>SECTION</th>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>REVIEW OF LITERATURE</td>
<td>3</td>
</tr>
<tr>
<td>III</td>
<td>PURPOSE OF THE STUDY</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Rationale of Test Selections</td>
<td>6</td>
</tr>
<tr>
<td>IV</td>
<td>METHOD</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Subjects</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Procedure and Materials</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Scoring the RBG</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Analysis of Data</td>
<td>16</td>
</tr>
<tr>
<td>V</td>
<td>RESULTS</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Table I Pearson Correlation Coefficients</td>
<td>19</td>
</tr>
<tr>
<td>VI</td>
<td>PROBLEMS FOR FUTURE STUDY</td>
<td>24</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
<td>29</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

In his address to the American Psychological Association convention of 1950, J. P. Guilford exhorted his colleagues to begin to turn their scientific attention to various aspects of creativity, areas long thought to be either unrelated to psychology or scientifically unobservable (1951). At that time, less than two percent of the books and articles published and indexed in the Psychological Abstracts were concerned with the area of creativity. By 1965, 310 such books and articles were listed, and the number is increasing each year (Wode, 1968). While research in the area of creativity has mostly been involved with the relationship of creativity to intelligence, more recently creativity has been studied as a separate entity.

In creativity, as perhaps in most areas of psychological research, there have been many differences of opinion and contradictory results. As Yamamoto (1967) points out, generally, investigators have not agreed on what would be the most accurate yet easily obtained index of creativity. Most of the more easily obtained measures have obvious shortcomings (examples include grade point average, teacher ratings, supervisor opinions, etc.) and there is a definite paucity of validation studies. Barron (1973)
found that even instructors at an art school were quite unreliable as judges of the creative ability and production of their students. Thus, we can see a need for objective measures in creativity research.
CHAPTER II

REVIEW OF THE LITERATURE

Much of the controversy over results obtained by the various methods used to discern creativity has centered on the relationship of creativity to intelligence. In their research with school children, Wallach and Kogan (1965a, b) found discrepancies between their results and the traditionally held view that a strong positive correlation exists between creativity and intelligence. However, their studies, which form a basis for much of the dissent concerning the relationship between creativity and intelligence, are experimentally weak in many ways. For example, by using only extremely gifted children (minimum IQ was 130 for their subjects) for these studies, they cut off the lower range of intelligence, which in this case included the majority in a normal population. Getzels and Jackson (1962) made the same error in their research on the gifted child, and not unexpectedly, found similar results, mainly, that there is a "ceiling effect" which occurs in the relationship of IQ to creativity up to a certain point and then found an inverse correlation from that point up. While these data are indisputable, the generalizations made by the authors of these studies are not. One cannot generalize findings from
a select, non-random sample to the general population.

Schubert (1973) explains these results as indicating that intelligence allows for development of creativity but does not in and of itself insure its development. Barron (1969) hypothesizes that the most probable state of affairs is that a low positive correlation exists between general intelligence and creativity. Individuals of varying degrees of creativity in professions intrinsically creative are of quite high measured intelligence, but their degree of creativity does not co-vary significantly with their intelligence scores.

Several studies have attempted to correlate different creativity tests. For example, Hargreaves and Bolton (1972) examined the relationship between 15 divergent and non-divergent tests of creativity, with two issues in mind. They wished to determine whether creativity is a unitary dimension across and within tests, and how this range of abilities is related to IQ. Their results led them to conclude that an integrated range of functions represented by the divergent tests were related to IQ in subjects of average ability, but these functions remained factorially distinct. One of the creativity tests used in this study was the Remote Association Test (RAT), developed by Mednick (1962) from his associational theory of creativity. In this study, the RAT correlated highly with both IQ and the divergent thinking tests.
There is a more pervasive problem in the area of creativity research, namely, how does one measure creativity? Nichols (1972) points out two major approaches to dealing with the concept of creativity in normal populations. One can either rate products on a continuum of creativity, and examine the personal and social factors associated with creative achievement, or one can identify certain people as creative on the basis of some criteria and explore the psychological significance of common personality traits in terms of their contributions to creative production. The divergent production tests, as exemplified by Guilford's Unusual Uses Test (1952), are the most commonly used tests of creativity, and are of the first approach.

Nichols (1972) goes on to say that approaches to creativity anchored to creative production are preferable to the trait based approach because the indices selected usually lend themselves more to operational definition of creative products. For example, Barron (1969) pointed out the shortcomings of the trait based approach of identifying individuals as highly creative and using them in creativity research (i.e. the subjectivity of GPA, teacher ratings, number of books published, etc.). Treffinger and Poggio (1972) call for more adequate conceptual and operational definitions in creativity research, and stress the importance of development of criteria for new measures of creative talent.
CHAPTER III

PURPOSE OF THE STUDY

The purpose of the present study was to modify an existing psychological instrument for use in creativity research and study its applicability. Briefly, the design was a simple one, that of giving a group of subjects three tests, two of which are valid and reliable tests of creative production and one which was of an experimentally developed test. Performances on the three tests could then be compared, and conclusions regarding concurrent validity drawn.

The creative production mode was chosen because of its applicability to operational and conceptual definitions of creativity and because of its objectivity as compared with the alternative method of testing and comparing personality traits of selected "creative" people.

RATIONALE OF TEST SELECTIONS

The test selected for modification and study was the Bender-Gestalt Test. The Bender-Gestalt Test consists of presenting nine stimulus cards consecutively to a subject or group of subjects. Each card has a design on it of one or more geometric figures or dots in various arrangements. The subject is asked to draw each of these designs. The only change in the presentation method in this study
was the instruction to the subject, "Make them so that they will be most pleasing to you."

Although the Bender-Gestalt has not been used in the past for creativity research, it has been shown to be an effective test for measuring other aspects of human behavior. It has been proven to be one of the most effective tests available for the diagnosis of diffuse neuropsychological impairment. Hutt (1969) has also indicated another use of the test, that of a projective instrument for personality assessment. Koppitz (1963) developed an additional scoring system for the Bender which has proven quite useful in discerning motor versus perceptual impairment, developmental disabilities, and developmental emotional problems in children.

The purpose behind selecting the RBG was the demand for a quick, easily and reliably scored test of creative production with a high degree of construct validity. Many of the currently available creativity tests and batteries are of limited utility. This is due to their lack of specific, operationally defined criteria, lengthy administration and scoring times, lack of cross-validation studies, and some very poor experimental techniques (Barron, 1969). The scoring system developed for this study to modify the Bender-Gestalt into a test of creative production entailed specific, reliable and objective criteria for determining the "creativity score" a subject could earn on a protocol.
Also, the combined administration and scoring time of 47 minutes is an improvement over many of the presently available tests of creativity or creative production.

The theory on which the Bender-Gestalt is based hypothesizes that the test taps the basic active, integrative process of visual perception and motor reproduction of the presented stimulus. The Bender-Gestalt is a quick and easy method of assessment with a satisfactorily high degree of reliability, and is, for the most part, independent of verbal communication. The precendents for manipulating and refining existing tests in order to measure other traits than those originally postulated have already been mentioned.

A revised version of the Bender-Gestalt Test (the RBG) was selected as the experimental test because it involved an active integrative component, perception of a stimulus and its motor reproduction. This is theoretically related to Mednick's (1967) associational theory of creative production in that, according to his model, the creator must also use active integrative processes. The term "active integrative" refers to the intellectual process of visually perceiving a stimulus, or a set of stimuli, and motorically reproducing the stimulus.

The two possible modifications that a subject could make in the RBG designs were simplification and complication of redrawing the presented stimulus designs. A subject could earn points in simplification by reducing the number
of figures and angles or by separating the figures spatially. Complication of the figure was rated according to an increase in the number of figures or angles, interpenetrations (i.e., joining figures together by overlapping them) and changes in shape or position. A complete explanation of the scoring system developed in this study may be found in the METHOD section of this thesis.

The second test used was Mednick's (1962) Remote Associations Test (RAT), which is a direct outgrowth of his associative theory of the creative process and production. It consists of three seemingly unrelated words which can be seen to relate to each other by the introjection of a fourth word. For example, the words "rat", "cottage" and "blue" can be seen to be related via the word "cheese."

The RAT uses statistical frequency as the basis for setting the criteria for a creative response. In well over ten validation studies carried out by Mednick (1962, 1967 and 1968) and others, the RAT was able to consistently differentiate high creative from low creative subjects. The RAT correlated significantly with faculty ratings (Mednick and Halpern, 1962), occupational interest on the Strong Vocational Interest Blank (Kowalski, 1961), grade point averages at MIT (Miller, 1960), originality and quantity of anagrams (Karp, 1960) and the Originality scale on the IPAR (Crutchfield, 1955). In Mednick's (1962) two published reliability studies the Spearman-Brown reliability of the RAT was .92 and .91.
The associative model of creative production is best explained by relating it to Guilford's (1969) Structure of Intellect model. Whereas Guilford distinguishes between two types of creative production, divergent (a search for logical alternatives) and convergent (a search for logical imperatives), Mednick's associative theory does not make this differentiation. To Mednick, the salient point is not classifying the type of production but determining the degree of difficulty of relating the problem to a solution. This solution may take the form of a logical imperative or a logical alternative. The more abstruse the association of a problem to a workable solution, the more "creative" the product. In behavioral terms, the less statistically frequent but logically possible a response is to a given stimulus, the more "creative" it is considered.

Mednick identifies this associative ability as creativity. To him, the creative process involves actively integrating a response or responses with a stimulus or stimuli. In this way, the associative theory of creative production can be seen to be related to the hypothesis underlying the Bender-Gestalt.

The third test for creativity used in this study was the Making Objects Test (MO), developed by Guilford (1953) and his associates as part of their divergent thinking battery. The MO consists of combining a given set of designs in such a way as to form a picture of a common object. For example, by using a quadrangle, a rectangle
and a circle, the subject is asked to make a picture of a lamp within a required amount of time. Guilford and Christian (1973) have found the MO to be a useful, expedient and valid test of non-verbal processes of creativity. Although the MO is usually used in conjunction with part or all of the Guilford Divergent Thinking Battery and thus has not been studied by itself, results of reliability and validity studies of the Battery indicate it to be an effective and reliable measure. The RAT and the MO are two tests of creativity, both verbal and non-verbal, that have not produced inconsistent and contradictory results and are amenable to group presentation methods.
CHAPTER IV

METHOD

SUBJECTS

Ninety students from a college introductory psychology class were used as subjects. Participation was partly voluntary in that Ss could elect either to take part in the study or to turn in a book report.

PROCEDURE AND MATERIALS

The Ss were given the three tests of creative production in one group meeting lasting approximately 55 minutes. The Remote Associations Test by Mednick was administered first, using the standard set of instructions which accompany the test. The test in its entirety would require 40 minutes to complete as many of the 30 test items as possible. Each item consists of three words which are related to each other in some way. For example, the words "book, tree, table" can be related to each other by the word "leaf". Since it was logistically impossible to test all the Ss in one session using the full length form of the RAT, only the odd numbered items were given, thus reducing the actual testing time limit to 20 minutes. This in no way affected the test results since both split half
and odd-even reliability (corrected for length) have been calculated at .92 and .91, respectively (Mednick, 1962). A S's score is the raw number of correct responses. A standard answer key is employed in scoring. In the event a response differs from the standard response but it conceivably related to the three stimulus words, credit is given if the association to all three stimulus words occurs less than one out of 1000 responses on the Kent-Rosanoff Word Association List (1910).

The Ss were next given the Making Objects Test according to the standard instruction set accompanying the test. The Ss were allowed to complete a sample item, then given three minutes to complete each of the two parts of the test. The test consists of Ss being instructed to draw nine figures in each part of the test using only combinations of six given geometrical figures. Ss must use only figures which are given at the top of the test protocol. A S's score is one point for each time a given figure is introduced in a different way in a part of the test.

Finally, the Revised Bender-Gestalt Test (RBG) was presented. The group presentation method used was similar to that of Koppitz (1963) in her work with brain damaged and emotionally disturbed children. Each of the nine standard Bender stimulus designs was projected upon a screen at the front of the testing room, accompanied by the single verbal instruction, "Make them so that they will be most pleasing to you." Each figure was presented for two minutes. At the
end of that time, a new figure was immediately presented. After the completion of the test, the S's were given a short debriefing session about the study.

SCORING THE RBG

The scoring of the RBG deserves some detail here (see Appendix). The RBG produces two sub-scores, a complication score (BGC) and a simplification score (BGS). The BGC is calculated by allowing a S one point for an increase in the number of figures, interpenetrations, number of angles in a figure, change of shape in a design as long as the original concept of the design is apparent, or a change (rotation) in the position of a design. A S scores one point for each of these five classes of complication of design within a single design. If at least one complication occurs within a single figure according to the above criteria, a scoring weight of four is added to the total score for that figure. For example, if a S adds two figures to the original design and rotates the design 180 degrees, he would score one point for each of these complications and also earn a score weight of four for the figure, giving him a total of six points for that particular design. The scoring weight is based upon the premise that in most instances a production with these figures involves more integrative and associative effort and skill than simplification. The S must draw more upon his imagination and associations from his past experiences and present mental function-
ing. Many examples of creativity enhanced by complexity can be found in art, from Durer's woodcuts to Dali's surrealistic paintings.

The second sub-score (BGS) consists of scoring one point for a reduction in the number of figures, number of angles in a figure or a separating of figures without any other changes. One point is scored for each of these criteria within a design, with no basal rating. The rationale behind this scoring can be related to artistic expression involving simplicity. Examples of this type of creative production based on paucity of detail can be found in Oriental drawing and painting, haiku poetry and some pieces of modern sculpture, where simplicity with emotional impact is considered aesthetically preferable and creative. The rationale behind not assigning a basal weighting to a simplification score lies in the problem of not being able to create the same amount of emotional impact (here defined as attentional drawing quality) with simplification as with complication, due to the paucity of elaboration.

A subject's score cannot reflect a simplification and complication score on the same figure. The decision to assign a simplification or complication score to a figure is based upon which group of modification criteria the figure most closely resembles, as outlined in the Appendix. This restriction is necessary because the two concepts of simplification and complication are theoretically mutually exclusive. Although in reality there are a few instances where
these two classes of modification could be present within a single design, the restriction is applied here for scoring ease.

Each S's total complication and simplification scores were kept separately for final analysis.

ANALYSIS OF DATA

Seven of the 90 S's scores were initially discarded because English had not been their first language. This was justified for reasons of experimental design, since the RAT was standardized on only native-born American students and adults. Also, the RAT requires a relatively close acquaintance with English and American-English colloquialisms.

The scores from the three tests were tabulated according to the aforementioned procedures and presented in table using a S by individual score method. The BGC and BGS for each score was totalled over all nine Bender designs, with each score being kept separate.

In addition, eight judges were used in this study. Four judges not already acquainted with the RBG scoring system were each given a brief written description of how to score the Bender protocols and scored a random sample of 10 tests. Interrater reliability was computed using the Spearman-Brown method. Four other judges were asked to give a holistic assessment of the sample protocols. No guidelines were given these judges, and they were instructed to
use any criteria they wished to rate the protocols on a five point discrete scale, with minus two being the least creative a protocol could be and plus two being the most creative. Interrater reliability was computed among the holistic judges using the same method as was used with the RBG instructed judges.

Finally, Pearson's Product Moment Correlation was computed for the RAT, MO, BGC, BGS, and the one holistic judge who was considerate enough to score all 83 test protocols. This comparison yielded a 5x5 correlation matrix. All data were coded and analyzed using the Harris SPSS Program (Statistical Package for Social Scientists). The program used computes the correlation matrix and tests the significance of each correlation.
CHAPTER V

RESULTS

Five significant correlation coefficients were discovered. The RAT correlated significantly with the BGC, the MO and the single judge's holistic assessment. All of these correlations were in the positive direction. The BGC was also found to correlate positively and significantly with the holistic assessment. Finally, the BGS and BGC were found to correlate significantly in a negative direction. This correlation matrix can be found in Table I.

Interrater reliability was found to be .89 for the RGB scoring system, which was satisfactory. Interrater reliability for the holistic assessment was computed as .56, which is not significant.
TABLE I

PEARSON CORRELATION COEFFICIENT

<table>
<thead>
<tr>
<th></th>
<th>RAT</th>
<th>BGC</th>
<th>BGS</th>
<th>MO</th>
<th>HOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAT</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S=0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGC</td>
<td>0.1935</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(83)</td>
<td>(0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S=0.040</td>
<td>S=0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGS</td>
<td>-0.0360</td>
<td>-0.5869</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(83)</td>
<td>(83)</td>
<td>(0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S=0.373</td>
<td>S=0.001</td>
<td>S=0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MO</td>
<td>0.2725</td>
<td>0.1035</td>
<td>0.0001</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(83)</td>
<td>(83)</td>
<td>(83)</td>
<td>(0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S=0.006</td>
<td>S=0.176</td>
<td>S=0.500</td>
<td>S=0.001</td>
<td></td>
</tr>
<tr>
<td>HOL</td>
<td>0.2253</td>
<td>0.2765</td>
<td>-0.0405</td>
<td>0.1335</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>(83)</td>
<td>(83)</td>
<td>(83)</td>
<td>(83)</td>
<td>(0)</td>
</tr>
<tr>
<td></td>
<td>S=0.020</td>
<td>S=0.0006</td>
<td>S=0.358</td>
<td>S=0.114</td>
<td>S=0.001</td>
</tr>
</tbody>
</table>

(COEFFICIENT / (CASES) / SIGNIFICANCE)

The results of this study support the hypothesis that the RBG and RAT are similar measures of creative production. The interesting point here is that although the two tests are based on similar theories, the RAT is a verbal test while the RBG is non-verbal for the most part. We can say with certainty that there is a positive relationship between the two. This can be explained in that both tests tap active integrative processes which contribute to creative production.

The inverse relationship between the BGC and BGS is most likely simply an artifact of the scoring methods, which are nearly mutually exclusive. The lack of any even
nearly significant correlation between the BGS and the other measures suggests that the "simplicity as creativity" approach is not very promising, at least as an approach to accounting for variation in scores on commonly accepted tests of creativity.

The significant correlation between the RAT and the BGC and the holistic assessment indicates, at first glance, a quick and easy method of assessment exists, that of simply letting any group of people use some sort of intrinsically developed criteria of "creativity" to determine people with creative ability and people without. However, the lack of a significant correlation between the single holistic judge who scored all the protocols and the remaining holistic judges who scored the protocol samples merely adds to the considerable previous evidence indicating the subjective judgement cannot scientifically or reliably be used as a criterion for creative production research at this time. Nevertheless, some individual judges of a particular test performance may be highly effective in terms of correlations with other valid tests, and in terms of their own repeat or split-half reliability performance.

The RAT and MO tests correlate significantly (p < .01). Here again we have a verbal and a non-verbal test which appear to be measuring a single operation or set of operations. In a Guilfordian scheme, these tests could be
said to represent convergent thinking. However, the RAT requires one correct solution to a problem (with exceptions as previously noted) while the MO requires as many original solutions as possible. There seem to be at least two possibilities: that both tests are in actuality measuring two aspects of the same operation; or, they are measuring two separate but related operations. However, from this study, one can only say with certainty that the two tests are correlated and appear to measure a common mediating construct.

The MO and BGC, interestingly enough, do not correlate significantly. This seems inconsistent with the other significant results obtained. It is particularly interesting to note the similarities between the two tests, i.e., both are basically non-verbal, figure reproduction tests. However, the MO scoring is "stricter" in that the examinee must make his figure at least resemble the required design, and is given only a limited number of designs which he may employ. The BGC, on the other hand, has absolutely no restrictions given at testing time.

The two tests, if viewed from the Guilfordian Structure of Intellect model approach can be seen as examples of convergent (MO) and divergent (RBG) thinking. Indeed, Guilford would probably explain the lack of a significant correlation between the two tests as indicative of the true separations of the two cognitive operations.
The associative explanation could be that, while both measure an integrative operation, the MO is subject to a finer quantification process, i.e., a S can earn points based on the degree of his figural manipulations on the MO test, while the S's score on the BGC has a tendency to reflect a more "all or none" approach to scoring. In the MO test, an S can earn as many points as he can by using the same figures in different ways. In the RBG (BGC), there is a maximum number of points that can be earned, one point for each class of manipulation.

The RBG has been shown, as proposed, to be an accurate, reliable measure of the active, integrative process which has, in this study, been called "creative" production. The Bender-Gestalt may become the test of choice because of its ability to simultaneously test for creative production ability, personality assessment and neuropsychiatric impairment. Judging from the strength of correlation, the MO test appears to also be a quick, easily scored test of creative production. However, the RBG has, in the writer's opinion, much more room for development and fine tuning, provided adequate research is done. The strong theoretical relationship between the RAT and RBG, along with their significant correlation, indicates both tests give the associative approach to creativity research some foundation.

One point worth discussing is the value of this type of research. Because someone scores well on some arbitrary
psychological test of "creative production", does this mean the person will be successful in an intrinsically creative field? Are there practical applications in the psychometry of "creativity" tests?

Any information a counselor, vocational or otherwise, can use in guiding an individual is potentially useful. From a consumerist point of view, the quickest yet most valid and reliable psychological tool (i.e., the most efficient psychometric instrument) is the most valuable. This type of research may be quite a distance from actual practical application, but it is the most scientific and logical approach to initial study of a problem that we have at this time. It can be suggested that the RBG is amenable to increased refinement and predictive efficiency and thus, potentially a useful, applied approach in the selection of individuals for creative performance in arts, graphic design or any field requiring imaginative fluency.
CHAPTER VI

PROBLEMS FOR FUTURE STUDY

This study and its results indicate some future avenues of research into the question of determining individuals with creative potential. The MO test, because of its significant correlation with the RAT, although coming out of a different theoretical school, deserves further study. Because of its short examination time and ease of scoring, the test is attractive to the psychometrician, providing it gains more credence from further validity research.

The RBG may be compared with other tests of creative production or tests of creativity based on the trait approach. An interesting study would be the comparison of a group of highly creative individuals selected by external career and their performance on the RBG with a group of Ss identified as creative only by criteria, with a group of individuals selected on the basis of extremely high BGC scores, with basic demographic variables controlled.

One might test the effects of stress on creative production by using a randomly selected group, split into control and experimental sections, and introducing a stress producing variable. After this procedure, both groups
would be tested on the RBG. A subject by treatments design analysis could then provide a further test of the RBG's efficacy.

Also worth consideration would be studies of the RBG correlations with various personality traits; neuroses, needs systems (i.e., aspiration level), organic brain damage, etc.

Within the next 20 years, psychology, as all other human endeavors, will continue to grow in scope and complexity. In order to continue this advancement, the cycle of research, validation or rejection, and more research will by necessity continue. Only in this manner can some objectivity be gained.
REFERENCES


Guilford, J. P. Reports from the University of Southern California. 1952, no. 7, USC Press.


APPENDIX

SCORING INSTRUCTIONS FOR THE REVISED BENDER-GESTALT TEST

These tests consist of nine designs or groups of designs. Although they may seem incomprehensible to you, the scoring of the tests is actually quite simple and straightforward.

There are two types of changes we are looking for in these tests. First, look at the original designs on the cards in front of you. Go through them one at a time to familiarize yourself with them. Now, place your first test sheet in front of you. There should be a number next to each design or group of designs. This number corresponds to each card.

There are two possible scores for each test. Only one score can apply to a single numbered design. The first score is called a complication score, and is arrived at by judging the design on the following criteria:

1. An increase in the number of figures from the original design on the card.
2. An increase in the number of interpenetrations from the original. Interpenetration means an overlap or a point where two figures touch.
3. An increase in the number of angles in a figure.
4. A change in the shape of the design.
5. A change in the position of the design. This could be a rotation of one or more of the figures in the design, or reversing the relative positions of the figures, as is possible in the case of design A.

If one of these types of changes occur, then score one point for that design. If two types of changes occur, then score two points. The maximum possible number of points in this type of scoring is, thus, five. If you have scored at least one point for this class of changes, add four more points to the person's score for that figure, and record this under the "COMP" column.

The other possible type of score for a test is called a simplification score. It is arrived at by scoring a test design on the following criteria:

1. A reduction in the number of figures from the original design.
2. A reduction in the number of angles in a figure.
3. A separation of the figures in a design (when there are more than one figure) without any other changes.

Score one point for each of these types of changes, and record this score under the "SIMP" column, next to the first column headed "COMP".
SCORING HINTS

A single design (that is, one of the nine numbered designs) can be scored only with a simplification or complication score, and only one of these types of scores for a design is allowed. If a design seems to have elements of both types of changes, the complication score takes precedence over the simplification.

In the case of the figures which contain circles or dots, do not count the individual circles or dots to determine if there is an increase or decrease in the number of figures. Score for increase in number of figures only if, in this case, there are other types of figures added to the original design.

Score zero if no changes occur from the original design, or if the drawing is so unrecognizable that you cannot see any of the elements from the original in the new drawing.

Remember that this is not an art contest, and do not rate the tests on the person's artistic ability. Try to adhere to the scoring system as closely as possible.