A New Method of Scoring
Torrance's Test of Creativity

By

Yalcin Ilsever

A THESIS IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Educational and Counselling Psychology, and Special Education: Measurement and Evaluation)

We accept this thesis as conforming to the required standard

UNIVERSITY OF BRITISH COLUMBIA

April 2000

Yalcin Ilsever, © 2000
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Education (MERI)

The University of British Columbia
Vancouver, Canada

Date April 17, 1990

(Signature)

Date /k. /l^ (ftUfm)

My fu- 1+

DE-6 (2/88)
ABSTRACT

This study proposed an alternative method of scoring the Torrance Test of Creative Thinking: Thinking Creatively with Words (TTCT: Verbal A). This scoring method was compared to the conventional scoring method. I first considered the current method of scoring, "the old method" and then proposed a "new method". The old and the new methods of scoring were compared on a subscale of the TTCT. Both the old and the new methods of scoring were then compared to the Canadian Test of Basic Skills (CTBS) total language scores to determine if a relationship existed between verbal academic achievement versus the creativity construct.

The new scoring method of creativity showed an improvement over the old method of scoring on the TTCT: Verbal A subtest. The reliability coefficient, for internal consistency, declined from 0.83 to 0.54. This decline was interpreted as an improvement because the initial value of 0.83 was artificially high. This was due to a very high intercorrelation between fluency and originality (inter-attribute
correlation value = .97). The new scoring method corrected this inflation.

I found a very low, negative correlation between CTBS total language and both the old and the new methods of scoring for creativity. The correlation coefficients were -.18, and -.17, respectively. This low value negative correlation demonstrated that creativity was not connected to verbal academic achievement.

Although divergent thinking is the most commonly used aspect of creative thinking, previous researchers (e.g., Runco, 1986) have suggested that the tests of creativity lack reliability. The issue of reliability of a test score has generally been defined in terms of variation of scores obtained by the individual on successive independent testing (Cronbach, 1947). Neither the assumption of constancy of true scores nor the assumptions of experimental independence can be realized in practice with most psychological variables. Therefore, the composite reliability of a test needs to be considered as a concept, which cannot be directly measured. Torrance (1986) specifically discouraged the use of composite test
scores and instead recommended interpretation of the subscale scores in relation to one another. The fundamental problem currently faced with the TTCT is that scores derived from the same response data cause potentially spuriously subscale correlation. This inflates the reliability coefficient and seriously jeopardizes score consistency and reliability.

Thorndike (1972) has argued that fluency, flexibility, and originality scores tended to be highly correlated, since all are accumulated over the same set of responses given by the examinees. This seriously discounts the value of the old scoring formula recommended by the Technical Scoring Formula of the TTCT. This fundamental problem of scoring has been attributed to derivation of five separate scores in that the scores were derived from the same response data (Heasusler & Thompson, 1988).

This type of score derivation introduced spuriously high scale correlation because the flexibility, fluency, and originality scores tended to be highly correlated.

Through some corrective procedures, correcting fluency, Clark
and Mirels (1970) recommended that correlation coefficient value from several measures of Torrance's Test of Creative Thinking (Figure Completion) would decrease. This procedure consisted of administering a revised form of the sub-test to 93 students. These were then scored, based on fluency, flexibility, originality, and elaboration. Corrected fluency scores led to lower reliability coefficient values. My study had a similar experience. The alpha coefficient value dropped to .54 from .83, with the new scoring method.

My research dealt with the score inflation, with the old scoring method, scoring inconsistency, and proposed a new method of scoring the TTCT. I also highlighted the issue of inter-attribute high correlation between fluency and originality attributes. These two attributes are the main ingredients in assessing creativity.

My research focused on the three traits of creativity, namely: (1) the fluency, (2) the originality, (3) the flexibility. These were measured with the use of the TTCT: Verbal A: Activity 4: Product Improvement. The TTCT was administered to a sample of children
(N=187) in grades 4 to 8. The data was analyzed to assess score reliability. The existing CTBS total language scores available at the school were also utilized in order to consider any relationship between creativity and verbal academic achievement.

The new scoring method enhanced the reliability coefficient and improved upon the scoring techniques. The new scoring approach adopted the same scoring techniques as the old method, with the measurement of fluency and flexibility attributes. However, a new approach was adopted. The new scoring method I have designed considered only the two score attributes: fluency and flexibility. The originality score, which is simply a "fall-out" from the fluency score, was not included. This exclusion avoided double counting between fluency and originality and dealt with the high correlation problem between these two attributes. With the old scoring method, the correlation between fluency and originality would be as high as .99 (Thorndike, 1972). This simply indicated that fluency and originality attributes were measuring the same creativity dimension. The new method of
scoring rectified this problem.

Hocevar (1979) recommended a method as a way of improving test reliability. Hocevar's method, however, recommended scoring for originality and flexibility and dividing the resulting score by the total number of responses. His recommendation was very valid in improving discriminant validity and did not address the score inflation and high inter-attribute correlation problem between fluency and originality. I have somewhat revised his recommended method and proposed that we capture the flexibility and fluency responses and divide the resulting score by the total number of responses. This ensured consistency in measurement and corrected the high inter-correlation problem between the TTCT attributes. It also improved score reliability.

Under the new method of scoring, each response for flexibility and fluency is adjusted by dividing the resulting scores by the total number of responses. Since fluency scores tended to be highly correlated with the originality scores, this new approach safeguarded against duplication in scoring also reduced spurious correlation.
between fluency and originality. Originality attribute was no longer required for scoring. Division of both flexibility and fluency by the total number of responses also streamlined the overall composite score.

The composite score of the TTCT, under the new method, contained fluency and flexibility scores. This also ensured that a number of responses which would normally be scored with "0", would be considered as part of the fluency score, as originality attribute was no longer considered.

The factor analytic results obtained by Heauslar and Thompson (1988) showed that the TTCT subscale yielded discrete scores. The general creativity factors were sufficient in assessing creative behavior. These factors included fluency, flexibility and originality. The total TTCT scores were not recommended (Torrance, 1974; Torrance & Ball, 1984) for assessment of creativity. He recommended using scores of fluency, flexibility and originality. This also jeopardized the integrity of the total creativity scores.

The new method improved the reliability of the test scores of
We TTCT by enhancing the existing scoring methodology. With the old scoring approach of originality, responses which did not meet the pre-determined fluency categories, were scored a "0", regardless of the relevancy of response. This diluted the originality scored, and distorted score reliability. The new method considered all the fluency responses and scrutinized all responses for consistency under fluency category. Fluency attribute was not simply considered as the number of total ideas, regardless of the degree of relevancy of task on hand. Each response of fluency was evaluated for its relevancy, in view of the task, on hand. This streamlined the scoring approach with fluency.

With the new method of scoring, the standardized alpha coefficient declined from .83 to .54. The new method of scoring improved the process and made it easy and more relevant to score for fluency and flexibility. A very low, negative linear relationship was found between the creativity construct and academic achievement, as represented through the CATS total language test.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF TABLES AND FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xv</td>
</tr>
<tr>
<td>CHAPTER ONE: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Introduction and the Nature of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>The problem</td>
<td>4</td>
</tr>
<tr>
<td>Benefits of the Study</td>
<td>17</td>
</tr>
<tr>
<td>Construct Definition</td>
<td>20</td>
</tr>
<tr>
<td>Scope of the Study</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER TWO: RELATED RESEARCH</td>
<td>25</td>
</tr>
<tr>
<td>Historical Theories and Definition of Creativity</td>
<td>26</td>
</tr>
<tr>
<td>Studies of Creativity in Children</td>
<td>34</td>
</tr>
<tr>
<td>Some Other Pertinent Studies with Children</td>
<td>35</td>
</tr>
<tr>
<td>The Hocevar Study</td>
<td>40</td>
</tr>
</tbody>
</table>
Further Studies of Creativity and Issues of Reliability..............44
Alpha Reliability Coefficient: Internal Score Consistency........48
Construct Validation Issues.............................................51
Tests of Creativity : Torrance...........................................54
Tests of Creativity: Guilford and Davis..............................57
The Hocevar and Michael Study of Validity..........................62
Score Reliability and Implications for Alpha Coefficient........65
Torrance's Test of Hypothesis: Interclass Reliability...............76
Summary.............................................................................80

CHAPTER THREE : METHODOLOGY...................................82
The Underlying Rationale for the TTCT...............................98
Mental Functions..................................................................85
Test Conditions.....................................................................87
The Scoring system.............................................................88
Methodological Issues in Measurement..................................94
Procedure and Measurement...............................................97
Procedure............................................................................99
LIST OF TABLES AND FIGURES

Table 5-1 Comparison of means for the two methods.............................111
Table 5-2 Comparison of means for the two methods of scoring, by grade Grade 4 Performance.................................................................112
Table 5-3 Comparison of means for the two methods of scoring, by grade Grade 5 Performance.................................................................116
Table 5-4 Comparison of means for the two methods of scoring, by grade Grade 6 Performance.................................................................117
Table 5-5 Comparison of means for the two methods of scoring, by grade Grade 7 performance.................................................................117
Table 5-6 Comparison of means for the two methods of scoring, by grade Grade 8 performance.................................................................118
Table 5-7 Correlation matrix Old Scoring Method..................................122
Table 5-8 Correlation matrix Old Method vs. New Method Total Score .........................................................................................................................124
Table 5-9 Item Total Statistics: Old Method of Scoring TTCT ..........137
Table 5-10 Item Total Statistics: Old Method of Scoring TTCT ..........138
Table 5-11 Item Total Statistics: New Method of Scoring TTCT ....139
Figure 1-1 Old method of measurement ........................................126
Figure 1-2 New method of measurement ........................................126
Figure 1-3 Total Mean Scores: Old Method vs. New Method
 ...........................................................................................................128
Figure 1-4 Total Mean Scores: The New and the Old Methods vs.
CTBC Total Language Mean Score (VB) ........................................130
ACKNOWLEDGMENTS

I am deeply grateful to my supervisor, Dr. Nand Kishor, for his exceptional guidance and for making himself available at all times throughout my research. My appreciation is also expressed to Dr. Marshall Arlin for his recommendations on the direction for my research and help he has provided in recommending the research topic. To Dr. Marion Porath, for her continued support and encouragement which have provided me with the privilege of being my research team and guiding me with her very creative questions and recommendations.