Is Creativity Unidimensional or Multidimensional?
Analyses of the Torrance Tests of Creative Thinking

Kyung Hee Kim
Eastern Michigan University

ABSTRACT: There is disagreement whether creativity is a unidimensional or multidimensional trait. The dimensionality of creativity is important to understand the mind’s cognitive functioning; thus aiding the development of human potential. Much of this dimensionality debate is related to the Torrance Tests of Creative Thinking (TTCT). Confirmatory factor analyses were thus conducted with data from 500 Grade-6 students, and several factor models were tested. The findings of this study show that the TTCT consists of 2 factors rather than a single factor, contrary to the majority of research on this subject.

The Torrance Tests of Creative Thinking (TTCT)

The TTCT was first published by E. Paul Torrance in 1966. There are two forms of the verbal and two forms of the figural. The tests have been renamed four times. The original purposes of the tests were for understanding strengths of students, research and experimentation, and general use for instructional planning (Torrance, 1966, 1974). Therefore, uses should reach for inclusion of higher scoring students, rather than exclusion of lower scoring students, for individualizing instructional programs (Treffinger, 1985).

The TTCT has been translated into over 35 different languages (Millar, 2002), and it is the most widely used test of creativity (G. A. Davis, 1997), including those for research purposes (Lissitz & Willhoft, 1985). The TTCT has been especially useful for identifying gifted and talented students and for counseling purposes (e.g., D. H. Cropley & Cropley, 2000). With 25 years of extensive development and evaluation (Millar, 2002), the TTCT has large norming samples, longitudinal validations (G. A. Davis, 1997), and high predictive validity for a broad age range (A. J. Cropley, 2000). The Figural form, especially, has equity benefits in terms of gender and race and for persons who have various language, socioeconomic status, and cultural backgrounds (Cramond, 1993; Torrance, 1977).

Content of the TTCT

Although there have been several revisions of the manual, the TTCT–Figural has remained relatively unchanged. The original edition in 1966 measured fluency, flexibility, originality, and elaboration, which were taken from Guilford’s divergent-thinking factors (Guilford, 1959; Torrance, 1966) and were only a little different from these subscales (fluency, originality, elaboration, abstractness of titles, resistance to premature closure, and creative strengths). The TTCT–Figural consists of three activities: picture construction, picture completion, and repeated figures of lines or circles; with 10 min to complete each activity (Torrance, 1966, 1974, 1984, 1990, 1998). In the latest version of the TTCT (Torrance, 1998), there are six different subscores; fluency, originality, elaboration, abstractness of titles, resistance to premature closure, and creative strengths, which are derived from the same response data. The raw subscales vary in range from 1 to 6 (for elaboration) to 0 to no upper limit (for fluency & originality). Therefore, standard scores for those subscales were used for this


Correspondence and requests for reprints should be sent to Kyung Hee Kim, 313k Porter Bldg., Teacher Education, Eastern Michigan University, Ypsilanti, MI 48197. E-mail: kkim7@emich.edu
study. The standard scores of each of the five variables are used according to the TTCT Norms-Technical Manual (Torrance, 1998) to get a “creativity index” (CI). Raw scores are converted into standard scores with means of 100 and standard deviations of 20. The subscale standard scores ranged as follows: fluency, 40 to 149; originality, 40 to 154; elaboration, 40 to 160; abstractness of titles, 40 to 160; resistance to premature closure, 40 to 160. The standard scores for each of the five norm-referenced measures are averaged to produce an overall indicator of creative potential. For the frequency of creative strength, a “+” or “++” is awarded on the basis of the scoring guide. The number of pluses are added (range for creative strengths: 0–26) to the averaged standard scores to yield a creative index (Torrance, 1984, 1990, 1998). Confirmatory factor analyses with creative strengths were compared with those without creative strengths for this study because the scoring procedure of creative strengths is different from the other five subscales.

Reliability

The TTCT–Figural manual of 1998 has provided the internal consistency reliability from the Kuder-Richardson 21 (KR–21) estimates. The reliability estimates of the creative index from the KR–21, using 99th percentile scores as the estimates of the number of items, ranged from .89 to .94. According to the TTCT–Figural manual of 1990, the intrarater reliability was above .90 (Torrance, 1990). According to the TTCT manuals of 1966 and 1974, the test–retest reliability coefficients (1-week, 2-week, 10-week, 6-month, and 3-year intervals) have ranged from .50 to .93, which is not so high. Treffinger (1985), however, concluded that given the complexity of creative thinking, the TTCT–Figural can be seen as having reasonable reliability.

Validity

There have been numerous validity studies involving the TTCT. Predictive validity studies for the TTCT scores have shown significant correlation with creative achievement in 7-year, 12-year, 22-year, and 40-year longitudinal studies (Torrance, 1969, 1980, 2002; Torrance & Wu, 1981). This is contrary to Baer’s (1994) opinion that the TTCT scores would not provide more information than intelligence test scores. According to the 40-year longitudinal study, 101 participants from the sample of 391 students in the original study from 1958 to 1964 had a high correlation (.45 for boys and .41 for girls) between creativity scores in elementary school and quality of creative production 40 years later (Torrance, 2002). The conclusion of the 40-year longitudinal study was that the total CI was a highly significant predictor for quality of creative achievement (Torrance, 2002).

Construct Evidence of the TTCT

Validity studies relevant to creativity measurement have grown from an extensive interest in tests of creativity. Analysis of the latent structure of an instrument is one useful way of examining a test’s construct validity. In this study, I analyzed the TTCT to understand its latent structure and to learn more about the cognitive function of creativity.

Because Guilford (1959, 1962) viewed divergent thinking as multidimensional, many researchers have come to the conclusion that creativity consists of several psychological factors. Similarly, Torrance (1966, 1974) discouraged the use of composite TTCT scores. He warned that using a single score could be misleading because each subscale score has an independent meaning.

However, studies on the TTCT have shown contrary results (Chase, 1985; Clapham, 1998; Dixon, 1979; Heusler & Thompson, 1988; Hocevar, 1979a, 1979b; Hocevar & Michael, 1979; Runco & Marz, 1992; Treffinger, 1985). Hocevar (1979a, 1979b) concluded that the TTCT and Guilford’s (1959) divergent thinking tests measured only one dimension rather than several independent dimensions. Dixon (1979) and Abernathy Tannehill (1998) noted a significant correlation between fluency and originality, and suggested that the subscores of the TTCT may actually measure similar constructs. Similarly, Heusler and Thompson, Chase, and Hassan (1985) all believed that the subscore correlations were too high to measure distinct traits. Hocevar and Michael reported that the heterotrait–monomethod coefficients were too high to be multidimensional. Runco and Marz criticized the lack of discriminant validity of the TTCT and other divergent thinking tests. Finally, Treffinger warned that interpretations of TTCT subscores as if they were independent should be avoided.
There have only been a few published factor analysis studies on the TTCT. Heausler and Thompson (1988) and Clapham (1998) concluded that the scores of the TTCT primarily reflected one general factor. Clapham (1998) noted that resistance to premature closure explained the highest amount of the variance in the CI. A number of factor analyses have been done with divergent thinking tests other than the TTCT (Michael & Bachelor, 1992; Proctor & Burnett, 2004; Runco, Plucker, & Lim, 2002–2001).

Combining Creative Style and the TTCT

What is the best way to explain the possible multidimensionality of the TTCT, if it exists? Guilford’s (1956, 1968) structure of intellect suggested multidimensionality, but it is also possible that there are only two factors. These may be explained using Kirton’s (1976) theory. Kirton (1976) proposed that creativity was composed of a single dimension ranging from an innovative to an adaptive orientation; the range reflected a person’s approach to creativity, problem solving, and decision making (Puccio, Treffinger, & Talbot, 1995). However, these two types may be separate dimensions. Two types of people are found based on my experience as a TTCT scorer and other scorers’ experiences. One type produced quick and novel responses, thus doing better on fluency and originality; the other type gave detailed responses, which indicated greater depth of thought and did better on elaboration and abstractness of titles. According to Kirton (1987), innovators prefer to create change by threatening the paradigm, whereas adaptors prefer to create change by working within the existing paradigm. A distinction, however, is that Kirton (1976) focused primarily on creative style not level of creativity (Gelade, 2002), indicating that a relation does not necessarily exist between a person’s style and creativity level.

However, Isaksen and Puccio (1988) suggested that a distinction between creative style and the TTCT as a measure of creative level might not be as clear as asserted by Kirton (1976). Isaksen and Puccio, as well as Torrance and Horng (1980), found that innovators were significantly more fluent and more original. Puccio et al. (1995) also found that the innovators gravitate toward creativity that was original, transformational, and expressive, whereas the adaptors were linked to creative endeavors that were logical, adequate, and well-crafted. Therefore, it may be that the innovative and adaptive types of creativity by Kirton (1987) might explain the latent structure of the TTCT. There are several studies that also investigated the relation between Kirton’s (1976, 1978, 1987, 1989) creative style and psychological type (e.g., Carne & Kirton, 1982; Gryskiewicz & Tullar, 1995; Isaksen, Lauer, & Wilson, 2003; Jacobson, 1993).

This study examines the following research questions using the TTCT data obtained: (a) Is creativity as measured by the TTCT a one, two, or multiple dimensional construct? (b) If the TTCT is two-dimensional construct, is there evidence that those two dimensions are innovative and adaptive?

In particular, the purpose of this study was to test if the proposed two-factor structure would fit the observed covariance matrices in the sample. The hypothesis is embedded in the model depicted in Figure 1. All relations were hypothesized to be positive. According to the proposed two-factor model, factor innovative was loaded by fluency and originality; factor adaptive loaded by elaboration, abstractness of titles, and creative strengths; and both factor innovative and factor adaptive loaded by resistance to premature closure. The double loading by resistance to premature closure is consistent with Torrance’s (1984, 1990, 1998) theory that creative people would be able to keep their mind open long enough to make mental leaps, whereas less creative individuals tend to prematurely leap to conclusions.

Method

Participants

For this study, the TTCT–Figural Form A was used, with data from 500 Grade-6 students, which included 242 boys and 258 girls (ages 10–15). The data were obtained through the Scholastic Testing Service Company. No information of sampling including relevant demographics was reported. The Scholastic Testing Service Company does not collect any ethnic or demographic data from examinees to enhance anonymity.

Data Analysis

The values of the mean, standard deviation, skewness, and kurtosis of the six TTCT subscales are re-
ported in Table 1. Multivariate normality is a common assumption in structural equation modeling, and it means that first, all the univariate distributions are normal; second, the joint distributions of any combination of the variables are also normal; and third, all bivariate scatterplots are linear and homoscedastic (Kline, 1998). The values of univariate skewness and kurtosis were examined to see whether each variable was approximately normally distributed. No values of the skewness and kurtosis were greater than |2.0|. Values of the relative multivariate kurtosis index produced by Prelis 2.53 were 1.117 for all participants.

Data were screened for outliers using DeCarlo’s (1997) SPSS Macro. There was an outlier, but no corrective action was taken because there were no differences in results when the outliers were removed. There were no missing values so that all of the 500 participants could be used for the analyses.

Confirmatory Factor Analyses

Confirmatory factor analyses were conducted to test the fit of several factor models including one-, two-, and, three-factor models using the LISREL 8.53 program (Jöreskog & Sörbom, 2002). The factor models with and without creative strengths were also analyzed. This was necessary because creative strengths had different procedures from the other subscales in scoring. Table 2 contains correlation matrices. All of the correlation coefficients between the variables were significant at the .01 level of significance except the correlations between creative strengths and resistance to premature closure (significant at the .05 alpha level), and between creative strengths and fluency (not significant at the .05 alpha level). In particular, the correlation coefficients between fluency and originality were very high (.844).

Finally, a confirmatory factor analysis with one general factor was conducted because several researchERS...
ers have concluded that the TTCT had only one general latent factor (e.g., Chase, 1985; Clapham, 1998; Dixon, 1979; Heausler & Thompson, 1988; Hocevar, 1979a, 1979b; Hocevar & Michael, 1979; Runco & Marz, 1992; Treffinger, 1985). Maximum likelihood (ML) is the most popular estimation method, and is the default in LISREL. ML has more accurate estimation of model fit and produces the best estimates of correct parameters when a model is misspecified. In addition, ML assumes multivariate normality, and the data were multivariate normal; therefore, ML was used as an estimation method.

**Results**

Fit indexes for the one-, two-, and three-factor models are shown in Table 3. In assessing model fit, I followed the two-index strategy and suggested indexes of Hu and Bentler (1998, 1999). This included reporting root mean square error of approximation (RMSEA) or standardized root mean square residual (SRMR) and supplementing it with one of the following: nonnormed fit index (NNFI; Tucker-Lewis index), incremental fit index (IFI), comparative fit index (CFI), or relative noncentrality index. Chi-square differences were computed to test the difference in fit among the one-, two-, and three-factor models. For the NNFI, IFI, and CFI, values vary between 0 and 1.0; values of .95 and above are considered to indicate a good model fit (Hu & Bentler, 1995, 1999). For the RMSEA (Steiger, 1990), values of about .05 are conventionally considered to indicate a close fit; values up to about .08 are considered reasonable, whereas Hu and Bentler (1999) recommended a cutoff close to .06. For the SRMR (Jöreskog & Sörbom, 1986), a cutoff of .08 or less is recommended by Hu and Bentler (1999). As Table 3 shows, the chi-square difference tests were significant, indicating that the two-factor model fits significantly better than a single-factor model, contrary to the majority of the research on this subject. This was supported by the values of the ad hoc fit indexes. Results indicate that the innovative factor is loaded by fluency, originality, and resistance to premature closure.

Table 2. *Correlations Between Subscales*

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Fluency</th>
<th>Originality</th>
<th>Elaboration</th>
<th>Titles</th>
<th>Closure</th>
<th>Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>.844**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Originality</td>
<td>.196**</td>
<td>.209**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaboration</td>
<td>.351**</td>
<td>.332**</td>
<td>.428**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Titles</td>
<td>.666**</td>
<td>.563**</td>
<td>.212**</td>
<td>.443**</td>
<td>.429**</td>
<td>.107*</td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strengths</td>
<td>.070</td>
<td>.134**</td>
<td>.573**</td>
<td>.429**</td>
<td>.107*</td>
<td></td>
</tr>
</tbody>
</table>

Note. Titles = abstractness of titles; closure = resistance to premature closure; strengths = creative strengths.
*p < .05, two-tailed. **p < .01.

Table 3. *Results of Model Comparison With One, Two, and Three Factors*

<table>
<thead>
<tr>
<th>Fit Index Group</th>
<th>Number of Factors</th>
<th>χ²</th>
<th>df</th>
<th>NNFI</th>
<th>IFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>Δχ²&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without Strengths</td>
<td>One</td>
<td>130.48**</td>
<td>5</td>
<td>.77</td>
<td>.89</td>
<td>.89</td>
<td>.22</td>
<td>.10</td>
<td>—</td>
</tr>
<tr>
<td>With Strengths</td>
<td>Two</td>
<td>7.82*</td>
<td>3</td>
<td>.99</td>
<td>1.00</td>
<td>1.00</td>
<td>.056</td>
<td>.021</td>
<td>122.66**</td>
</tr>
<tr>
<td></td>
<td>One</td>
<td>373.71**</td>
<td>9</td>
<td>.56</td>
<td>.74</td>
<td>.73</td>
<td>.28</td>
<td>.17</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Two</td>
<td>96.97**</td>
<td>7</td>
<td>.86</td>
<td>.93</td>
<td>.93</td>
<td>.16</td>
<td>.090</td>
<td>276.74**</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>Cannot be identified</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. NNFI = non-normed fit index; IFI = incremental fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; Strengths = creative strengths.
<sup>a</sup>Comparison is to the previous model.
*p < .05. **p < .01.
whereas the adaptive factor is loaded by elaboration, abstractness of titles, and creative strengths. However, the factor models without creative strengths fit better than those with creative strengths. This indicates that creative strengths might represent a separate factor. However, more indicators of the TTCT would be needed to test this model.

The two-factor model was further analyzed by examining parameter estimates. The large values of the factor loadings and large $R^2$ values indicated that the subscales were a good measure of their factors. However, the low $R^2$ value for elaboration (.21) suggested that this subscale is not as highly related to the adaptive factor as the other subscales loaded on that factor.

Post hoc modifications of the two-factor model were undertaken despite the practically nonsignificant indications of the results. The modification indexes (MIs) were examined to determine which subscales can parameter values differ for a better fit. Values of the MIs for parameters represent the amount that the overall chi-square value would decrease if the parameter were allowed to add or differ. These examinations revealed several factor loadings and measurement error covariances that would be added or changed. The MIs for the factor loadings indicated that if the values of the factor loading from adaptive to originality (decrease in chi-square: 5.19) groups had been allowed to add, the amounts of the overall chi-square values would have decreased, leading to improvements of the model fit. There were no measurement error covariances in our model; thus, none of the MIs referred to constraints on parameters. The MIs for the error covariances here indicated that if the error covariances between originality and resistance to premature closure were allowed to differ (decrease in chi-square: 28.17), the model fits could be improved. However, as we see, these decreases in chi-square are trivial, indicating the two-factor model fits very well.

**Discussion**

This study examined the possibility of a two-factor model based on Kirton’s (1976, 1978, 1987, 1989) adaptor–innovator theory. The hypothesized relations of the proposed model with both innovative and adaptive factors would be a good fit for the entire sample. According to this model, the innovative factor was loaded by fluency and originality; the adaptive factor

was loaded by elaboration and abstractness of titles; and both factors were loaded by resistance to premature closure. These relations between the factors and the five subscales were on the basis of Kirton’s (1987) descriptions of innovative and adaptive, Puccio et al’s (1995) findings, the scoring experiences of the TTCT, and Torrance and other researchers’ findings (Torrance & Horng, 1980; Isaksen & Puccio, 1988). The logic for the double loading by resistance to premature closure originated from Torrance’s (1984, 1990, 1998) theoretical assumption of creative individuals. Based on the results of this study, the proposed two-factor structure of the TTCT fits well.

Confirmatory factor analyses with one general factor were also conducted to compare the results with one factor to those with two hypothesized factors in this study. The chi-square values and different fit indexes suggested that the proposed two-factor model in this study was a much better fit. These findings are inconsistent with the empirical and theoretical literature on the TTCT in that Torrance (1966, 1974) suggested six different factors, and in that several factor analytic studies related to the TTCT concluded that the TTCT had only one factor. However, based on experiences in scoring the TTCT, the results can be considered as reasonable because the TTCT scoring trainers have found the same pattern of results from a long experience in scoring. In addition, the findings in this study were consistent with the descriptions of innovative and adaptive in Kirton’s (1987) adaptor–innovator theory and with other researchers’ findings (Puccio et al., 1995).

The finding that all of the correlation coefficients between the variables were high, particularly the correlation coefficient between fluency and originality (.844), is consistent with other research (Abernathy Tannehill, 1997; Chase, 1985; Dixon, 1979; Heausler & Thompson, 1988). This is also consistent with the work of Torrance and Safter (1999), who reasoned that the person who produces a large number of alternatives is more likely to produce original ones. The large measurement error variances between fluency and originality might also be explained by this reasoning. Simonton (1990) also found that a person’s originality is a function of the number of ideas formulated. In addition, the high correlations between all the variables here might partly result from the fact that the five different subscales were based on the same stimuli. Because of the high correlations among the subscales—especially the correlation between fluency and originality (from .79 and .86)—many

K. H. Kim
researchers (Chase, 1985; Clapham, 1998; Dixon, 1979; Heausler, & Thompson, 1988; Hocevar, 1979a, 1979b; Hocevar & Michael, 1979; Runco & Marz, 1992; Treffinger, 1985) insisted that the TTCT measures only one factor or that the subscales are not independent. However, the analyses conducted in this study support a two-factor structure in which the factors are correlated at .37. For example, correlations between some of the subscales on the fourth edition of the Stanford–Binet Intelligence Scale (Thorndike, Hagen, & Sattler, 1986) range from .66 and .73 (N = 5,000); those on the Woodcock Reading Mastery Tests (Woodcock, 1973) range from .77 and .89 (N = 103); those on the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983) range from .66 and .75 (N = 500 for preschool children and N = 1,500 for school-age children); and those on the Wechsler Intelligence Scale for Children–Third Edition (Wechsler, 1991) range between .69 and .70 (N = 2,200).

In conclusion, the results of the confirmatory factor analyses indicated that the two-factor model proposed in this study had a much better fit than the one general factor model. The TTCT consists of two factors rather than a single factor, contrary to the majority of the research on this subject. The innovative factor may consist of fluency, originality, and resistance to premature closure, whereas the adaptive factor may consist of elaboration, abstractness of titles, and creative strengths. The findings of this study confirm the observations of the test scorers and are consistent with several researchers’ (e.g., Isaksen & Puccio, 1988; Puccio et al., 1995; Torrance & Horng, 1980) interpretation of Kirton’s (1976, 1978, 1987, 1989) model. However, as there have been no published studies using creative strengths for analyses of the TTCT, because of the different procedures in scoring and other reasons as previously noted. However, this subscale is too important to be excluded from full explanations of the scores of the TTCT (E. P. Torrance, personal communication, October 30, 2002).

For future research, the direct relation between Kirton’s (1987) Innovative–Adaptive Inventory and the TTCT–Figural could be conducted using these analysis methods. Studies using item response theory would also be useful for future studies. Furthermore, in future studies the structure of subscores should be examined to study how each of the subscore predicts creative achievement (Clapham, 1998) and how it is related to each subscore and to the subscores of other divergent thinking tests.

**Limitations**

There are several limitations of this study. First, this study was based on only sixth-grade students. Therefore, the results of this study cannot be generalized for populations of all ages.

Second, the TTCT is a complex measure, and is complicated by the dependence of all scales on the same stimuli; as the high correlation coefficients between the subscales have shown in previous studies. Third, various studies have suggested different best predictors and number of factors. Therefore, from the disparity in the results of several factor analytic studies on the TTCT including this study, it can be hypothesized that the TTCT is measuring creativity differently for dissimilar groups. This would indicate that creativity may not exist in the same form for all groups. Therefore, multiple group analyses would be helpful to understand the latent structure of the TTCT. However, this study did not have a large enough sample size of diverse groups so that multiple group analyses could be run. Therefore, multiple group analyses should be conducted with several different groups for future studies. Fourth, the data were obtained through the Scholastic Testing Service Company, and no information regarding relevant demographics was provided. Therefore, we cannot generalize these results, and we should consider them as tentative until replicated with other samples. Fifth, many researchers have found that motivation (Bamber, 1973; Halpin & Halpin, 1973; Torrance,
1966, 1974) and testing conditions (Bamber, 1973; Callahan, 1991; Halpin & Halpin, 1973), as well as exposure to diverse information (Clapham, 2000–2001), influence TTCT–Figural scores. Therefore, race, socioeconomic status, and other environmental factors of participants should also be considered in confirmatory factor analyses or multiple group analyses using the TTCT, besides gender or grade differences, as analyzed in this study. This could give more information about understanding creativity tests including the TTCT; the nature of creativity; and, ultimately, how to encourage creativity in each individual.

References


Analyses of the TTCT


Creativity Research Journal 259
Copyright of Creativity Research Journal is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.