Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores

ABSTRACT

There is disagreement among researchers about whether IQ tests or divergent thinking (DT) tests are better predictors of creative achievement. Resolving this dispute is complicated by the fact that some research has shown a relationship between IQ and DT test scores (e.g., Runco & Albert, 1986; Wallach, 1970). The present study conducted meta-analyses of the relationships between creative achievement and both IQ and DT test scores. The analyses included 17 studies (with 5,544 participants) that established the correlation coefficients between IQ and creative achievement and 27 studies (with 47,197 participants) that established the correlation coefficients between DT test scores and creative achievement. Marginal, but statistically significant, Fisher’s Z-transformed correlation coefficients were revealed. The analysis found a significantly higher relationship between DT test scores and creative achievement ($r = .216$) than between IQ test scores and creative achievement ($r = .167$). The differences in the correlation coefficients were explained by differences in DT tests, creative achievement types, predicted time periods, and creativity subscales. The significant independent moderator effect for different DT tests indicates that the Torrance Tests of Creative Thinking (TTCT) predict creative achievement better than any other DT test included in this study. Among the creative achievement types, music is predicted the best by IQ and all others are predicted best by DT tests. Among the time periods evaluated, the relationship between DT test scores and creative achievement had the highest correlation at the period of 11-15 years.

INTRODUCTION

The value of any test is dependent upon its ability to predict performance in the characteristic being measured (Cramond, 1994). One often used example of a valuable test is the Binet intelligence test which lacks the capacity to measure more than a few human abilities and which was originally designed to identify
intellectual deficits. However, the Binet intelligence test has endured and even been updated to its current form, the Stanford-Binet Intelligence Scales. One of the reasons is that children with high scores on the Binet intelligence test tend to well in school.

One characteristic, other than IQ, that society has come to value over the past 50 years is creativity. The definition of “giftedness” is being updated to include individuals with high creative potential rather than just high IQ (e.g., Georgia Department of Education; Renzulli, 1986). Thus, measures of creativity are becoming increasingly important.

Guilford (1950, 1962, 1966, 1968) hypothesized that creative individuals possess divergent thinking abilities including idea production, fluency, flexibility, and originality. Guilford also argued that traditional intelligence tests (such as the Binet IQ test) do not measure some or all of these creative abilities. Guilford’s theories spawned an array of divergent thinking (DT) tests such as the Torrance Tests of Creative Thinking (TTCT), Wallach & Kogan Divergent Thinking Tasks, and Guilford Divergent Thinking Tasks. It also spawned research that attempts to correlate scores on these DT tests with creative potential.

Research has generally concluded that DT test scores and creativity inventory scores are predictive of creative activities, interests, and accomplishments later in life (Cline, Richards, & Needham, 1963; Kogan & Pankove, 1972; Russ, Robins, & Christiano, 1999; Torrance, 2002, Rimm & Davis, 1976; Wakefield, 1985). However, the rationale that DT is related to creative potential remains controversial. Plucker (1999) found previous research that suggested that DT test scores have little predictive validity with respect to adult creative achievement, but his own research tended to support the predictive validity of DT tests. Prior to Torrance’s longitudinal studies on his DT tests, the TTCT, skeptics expressed doubt that children that scored highly on the TTCT would produce useful creative achievement later in life (Torrance, 1972). Research has also indicated that the best predictor of creative achievement is past creative performance (Holland, 1961; McDermid, 1965; Taylor & Ellison, 1962) and that DT tests are no better than intelligence tests as predictors of real-life creativity (Cropley, 1972; Getzels & Csikszentmihalyi, 1964; Kogan & Pankove, 1974; Skager, Klein, & Schultz, 1967; Torrance, 1972; Wodtke, 1964). This controversy leads to a second issue: Do IQ tests or DT tests predict creative achievement better?

Researchers disagree whether DT tests measure significantly different traits than IQ tests. Some research has shown that divergent thinking tasks, DT test scores, and creative achievement are independent from IQ (e.g., Getzels & Jackson, 1958; Gough, 1976; Helson, 1971; Helson & Crutchfield 1970; Herr, Moore, & Hansen, 1965; Rossman & Horn, 1972; Rotter, Langland & Berger, 1971.

---

1 Some researchers (e.g., Torrance, 1984) refer to their tests as creativity tests. However, this article attempts to address this underlying issue (whether DT tests measure creativity) so for the purposes of this article the author has chosen to refer to DT tests.
Meta-analyses attempt to resolve apparent conflicts in literature; to discover consistencies and account variability in similarly appearing studies; and to identify core issues for future research (Cooper & Hedges, 1994). Thus, this study synthesized empirical research in the areas of creativity, DT tests, and intelligence tests for the purpose of creating a generalization about the relationships between creative achievement and both IQ and DT test scores. The five primary purposes of this synthesis were to:

1. Conduct a meta-analysis of correlations between IQ and creative achievement;
2. Conduct a meta-analysis of correlations between DT test scores and creative achievement;
3. Identify some of the variables that moderate those correlations and test whether correlation coefficient sizes vary systematically across differing levels of variables that are posited to influence the relationship between IQ and creative achievement (e.g., gender, IQ tests, creative achievement types, & predicted time periods²).
4. Identify some of the variables that moderate those correlations and test whether correlation coefficient sizes vary systematically across differing levels of variables that are posited to influence the relationship between DT test scores & creative achievement (e.g., gender, DT tests, creative achievement types, predicted time periods, & creativity subscales).
5. Use the correlations derived from the meta-analyses to investigate models of the relationships between creative achievement and both IQ and DT test scores.

² “Predicted time periods” refer to the future time period that was studied which was related to the results of an earlier test.
METHODS

Literature Searches

Over one hundred studies published from 1958 through the summer of 2005 were located in computer databases (e.g., Academic Search Premier, ERIC, Primary Search, PsycARTICLES, & PsycINFO) and bibliographic searches of each reference within the body of English creativity and intelligence literature. The keywords for the searches included: IQ, creativity, divergent thinking, intelligence, and creative achievement.

The criteria for inclusion in this meta analysis are: i) the reporting of the correlations between measures of intelligence and creative achievement, and/or the correlations between measures of DT and creative achievement; ii) a sample size for each correlation; iii) identification of the measure of IQ and/or DT; iv) and reporting of creative achievement in the area of quality or quantity, art, music, writing, science (including math, medical, and engineering), leadership, or social skills.

The above criteria were required by the limitations of statistical analysis. However, these criteria significantly limited the number of previous studies included in this meta-analysis. Many of the previous studies examined failed to report detailed information on procedures and/or results, making it difficult or impossible to generate correlation coefficients. Ideally, a meta-analysis researcher must try to collect and analyze the results of all existent studies related to the defined population. However, not all of the studies can be used because the statistics necessary for the analysis is not reported in some of them. According to Hunter et al. (1982), however, meta-analytic procedures are still valid even for convenience samples so that a complete sampling of the population is not necessary.

Where non-significant correlation coefficients were reported and interpretation was not available through sample size and probability estimates (e.g., the correlation coefficients between IQ [Terman Concept Mastery Test] and creative achievement in Hall, 1972; the correlation coefficients between IQ [Milta Group Intelligence test] and creative achievement in Milgram & Hong, 199), a Pearson correlation coefficient of \( r = .00 \) was assumed (Rosenthal, 1991). Some of the previous studies were not used because they compared t-tests between high and low creative individuals (e.g., Carroll & Howieson, 1992; Getzels & Jackson, 1961; 1962; Lynch, 1970); they presented comparisons between creative and control groups (e.g., Anastasi & Schaefer, 1971; Lang & Ryba, 1976); they reported multiple regression analysis for the prediction of IQ, creativity, leadership, and academic achievement (Chan, 2001) or the Scales for Rating the Behavioral Characteristics of Superior Students (SRBCSS, Chan, 2000); or they reported correlation coefficients within limited IQ ranges (Gonzales & Campos, 1997).

Ultimately, only 17 studies that presented correlation coefficients between IQ and Creative Achievement and 27 studies that presented correlation coefficients between DT test scores and creative achievement were included in the present meta-analyses.
Effect Size Calculations

Two quantitative syntheses of the resulting 17 studies and 27 studies were conducted, which was assisted by Schwarzer’s Meta 5.3 statistical software (1991). Fisher’s z transformation of $r$ was used for analyses to adjust for the non-normal distribution of $r$. Most researchers believe that studies which employ large samples should get more credit than those which are based on small samples because correlations become more stable as sample size increases (Schwarzer, 1991). Thus, the effect size $zr$ was weighted by sample size: the weighted mean $zr = \frac{\sum(N_j - 3) zr_j}{\sum(N_j - 3)}$, in this equation $zr_j$ is the Fisher $zr$, corresponding to any $r$ (Rosenthal, 1991). Each mean $r$ was tested using a random effects model of variance, and the reported values of $r$ are back transformations from $z$ (Hunter, Schmidt, & Jackson, 1982). A correlation coefficient was judged heterogeneous when the residual standard deviation exceeded ¾ of the effect size, and sampling error was less than 75% of the observed variance (Schwarzer, 1991). The residual standard deviation was used as the standard error in estimating 95% confidence intervals.

Ultimately, 368 correlation coefficients were retrieved from the studies: 94 correlation coefficients between IQ and Creative Achievement (from 17 studies involving a total sample size of 5,544 people); and 274 correlation coefficients between DT test scores and Creative Achievement (from 27 studies involving a total sample size of 47,197 people). Multiple correlation coefficients were obtained from studies that included more than one correlation using several creativity subscales or reported separate results for gender. A conservative statistical criterion ($p < .001$) was used to protect against Type I error (Rosenthal, 1991; Rosenthal & Rubin, 1984) when several correlation coefficients resulted in a study of moderators. Finally, as reported above, individual correlation coefficients were weighted by sample sizes (Hedges & Olkin, 1985; Hunter & Schmidt, 1990; Rosenthal, 1991) to adjust for sampling errors in $r$, giving more credence to studies with large samples because small sample sizes can increase Type II error (Hedges & Olkin, 1985).

Moderator Analyses

An analysis of moderator variables was conducted to test whether correlation coefficient sizes varied systematically because of the different variables that influence the relationships between IQ and Creative Achievement and between DT test scores and Creative Achievement. Correlation coefficient size heterogeneity can be explained through moderator variables. The moderator variables were categorized in a manner that seemed theoretically valid.

Previous research is inconclusive about whether gender affects creativity scores, thus a moderator analysis of gender was conducted to test whether correlation coefficient sizes vary systematically across male, female, and combined subjects. Previous studies found gender differences (e.g., Gupta, 1981; Jaquish & Ripple,
Whereas other studies (e.g., Ogawa, Kuehn-Ebert, & De Vito, 1991; Runco, 1991; Saeki, Fan, & Van Dusen, 2001) found no differences. Even among those studies that found gender differences there were inconsistencies in that some reported higher scores for males, others reported higher scores for females.

Previous research is inconsistent about whether DT test scores predict real-life creativity in a specific area of creativity such as art, math-science, drama, writing, music, etc. (e.g., Hocevar, 1980; Milgram & Milgram, 1976; Wallacch & Wing, 1969). Thus, a moderator analysis of Creative Achievement Type was conducted to test whether correlation coefficient sizes vary systematically in specific areas such as art, music, writing, science (including math, medial, and engineering), leadership, and social skills. However, some studies only presented results in general terms so the analysis included more general creativity areas, e.g., general creativity, quality of creative achievement, or quantity of creative achievement. Further, when data were entered, it was discovered that the studies varied how they classified specific areas, thus, for the present meta-analysis some of the fields were reclassified. For instance, in Carson, Peterson, and Higgins’ study (2003), “accomplishment in art and science fields” was classified as general creative achievement. In Baron’s study (1963), “overall performance in charades” was classified as quantity; “total effectiveness in improvisation” was classified as quality; both “overall effectiveness in command functions” and “overall effectiveness in staff function in the air force” were classified as quality. In Frederiksen and Ward’s study (1978), “number of activities” in the Test of Scientific Thinking was classified as quantity; “publications” were classified as writing; “collaborative research” was classified as science; “number” was classified as fluency; “unusual” was classified as originality; “teacher rating” or “professor rating” (e.g., Lunneborg & Lunneborg, 1968; Rimm & Davis, 1976), “supervisor rating” (e.g., Getzels & Csikszentmihalyi, 1976), and “teacher rating” and “principal rating” (e.g., Storm & Larimore, 1970) were classified as general creative achievement; however in Gough’s studies (1975, 1976), “supervisor rating” and “peer rating” were classified as science because their fields were science or engineering; “professors rating” was classified as art (Lunneborg & Lunneborg, 1968); and “business contact” was classified as social skills whereas “business organization” was classified as leadership (Wakefield, 1985).

Barron and Harrington’s comprehensive review of creativity research (1981) found that correlation coefficients between creative thinking abilities and intelligence vary widely depending upon the DT test being studied, the heterogeneity of the sample, and even testing conditions. Different results from different studies may also be due to different IQ tests and/or different DT tests. Thus, a moderator analyses of IQ Tests and DT Tests was conducted to test whether correlation coefficient sizes vary systematically between the Torrance Tests of Creative Thinking (TTCT), Guilford Divergent Thinking Tasks, Wallach & Kogan Divergent Thinking Tasks, Sounds and Images (see Khatena, 1971), Group Inventory for Finding
Meta-analyses of the Relationship


Different results from different studies may be due to different creativity subscales, thus, a moderator analyses of Creativity Subscale was conducted to test whether correlation coefficient sizes vary systematically across fluency, flexibility, originality, elaboration, inventiveness, creative strengths, and general creativity. Again when data was entered, it was discovered that the studies varied how they classified specific creativity subscales, thus, for the present meta-analysis some of the fields were reclassified: “associational fluency”, “ideational fluency”, “verbal fluency”, and “figural fluency” were classified as fluency; “spontaneous flexibility”, “adaptive flexibility”, “verbal flexibility”, and “figural flexibility” were classified as flexibility; and “verbal originality” and “figural originality” were categorized as originality. In several of the DT tests in the present study (e.g., Flescher, 1963; Hocevar, 1980; Kogan & Pankove, 1972; Vernon, 1972; Wallach & Kogan, 1965), subscales were not reported. Thus, the tasks in these studies were examined and categorized based on the nature of those tasks. In Cline, Richards, and Needham (1963), “consequences immediate”, “brick uses-total”, and “word association” were classified as general creativity subscale; “hidden figures”, “brick uses-change”, and “match problems” were classified as flexibility; “consequences-remote” was classified as originality. In Gough (1975, 1976), “total responses” were classified as fluency, and “Word Association Test” (Cline, Richards, & Needham, 1963; Gough, 1975, 1976) was classified as general creativity subscale. In Lunneborg and Lunneborg’s study (1968), “Architectural School Aptitude Test (ASAT)” was classified as general.

Different results from different studies may be due to the time period each study investigated, thus a moderator analysis of Predicted Time Period was conducted to test whether correlation coefficient sizes vary systematically from the present and in years 1-5, 6-10, 11-15, 16-20, and 21 and over.

Factors that might moderate the estimated population correlation coefficients between IQ and Creative Achievement were considered by comparing correlation coefficients based on Gender, IQ Test, Creative Achievement Type, and Predicted Time Period. Factors that might moderate the estimated population correlation coefficients between DT test scores and Creative Achievement in the present study were considered by comparing correlation coefficients based on Gender, Creative Achievement Type, DT Test, Creativity Subscale, and Predicted Time Period. Most studies did not report information about their subjects’ ethnicity so ethnicity was not considered as a moderator variable.

Hedges’ $g$ was derived for each individual correlation coefficient in the moderator analysis using Johnson’s DSTAT 1.10 meta-analysis software (1993). When significant moderator effects were found, the proportion of variance caused by the moderators was determined. Variables that were significant moderators were entered into a weighted linear multiple regression model (weighted by $N-3$, [Rosenthal, 1991]) using SPSS to clarify their independent effects for explaining variations in $zr$. 

112
RESULTS

A stem-and-leaf display of the 94 effects between IQ scores and Creative Achievement is presented in Figure 1, and a stem-and-leaf display of the 274 effects between DT test scores and Creative Achievement is presented in Figure 2. Both of the stem-and-leaf displays indicate a nearly normal distribution of correlation coefficients. After weighting by sample size, the mean value of \( r \) between IQ and Creative Achievement was .167 (95% CI = .141 – .193) and the mean value of \( r \) between DT test scores and Creative Achievement was .216 (95% CI = .207 – .225). Thus, a higher correlation coefficient was found between DT test scores and Creative Achievement than between IQ and Creative Achievement. Both of the mean correlation coefficients were judged to be small and statistically significant. However, both of the correlation coefficients were heterogeneous: Q (93) for correlation coefficients between IQ and Creative Achievement was 174.514 \((p < .0001)\), whereas Q (273) for correlation coefficients between DT test scores and Creative Achievement was 842.431 \((p < .0001)\); both of the residual standard deviations were greater than one fourth of population correlation coefficient size.

FIGURE 1. Stem-and-Leaf-Display for 94 Correlation Coefficients (r) between IQ and Creative Achievement.

- .9 I
- .8 I
- .7 I
- .6 I
- .5 I
- .4 I
- .3 I
- .2 I 24
- .1 I 122678
- .0 I 0000113345599
+ .0 I 0000244566889
+ .1 I 0012333334566677
+ .2 I 0111112224456677888999
+ .3 I 013355578888
+ .4 I 0014455566678
+ .5 I
+ .6 I
+ .7 I
+ .8 I
+ .9 I

Note. The correlation coefficients \( r \) ranged from -.24 in the first row of the display to +.48 in the last row.
Meta-analyses of the Relationship (.042 for the correlation coefficients between IQ and Creative Achievement; .054 for the correlation coefficients between DT test scores and Creative Achievement); Percent of observed variance accounted for by sampling error for the correlation coefficients between IQ and Creative Achievement was 52.81% and between DT test scores and Creative Achievement was 33.39%, and both of which were less than 75%. Because the data set was heterogeneous, it was necessary to search for moderators that may account for the remaining systematic variation in the data (see Schwarzer, 1991). The analyses of moderators were performed by breaking down the data into at least three subsets with respect to theoretically relevant variables. Moderator analyses were conducted to determine whether moderators describing subjects or features of IQ and Creative Achievement and DT test scores and Creative Achievement might account for variability in the magnitude of correlation coefficients. Each simple contrast tested as $z$ at $p < .001$. Correlation coefficients for each moderator are reported in Table 1 through Table 5 as mean $r$. Variability in the magnitude of correlation coefficients between IQ and Creative Achievement for Gender (QB [2] = 27.582, $p < .0001$), IQ Test (QB [3] = 102.415, $p < .0001$)

![FIGURE 2. Stem-and-Leaf-Display for 274 Correlation Coefficients (r) between DT test scores and Creative Achievement.](image_url)

| -9 | 1
| -8 | 1
| -7 | 1
| -6 | 1
| -5 | 1
| -4 | 1
| -3 | 00
| -2 | 13778
| -1 | 01224455778
| 0 | 0222233345558
| +0 | 0112334455678899
| +1 | 1222334455556678888999999
| +2 | 00001111122222222344555555555566777777788888888999999
| +3 | 0000000111111222222223344444444445555555555566666667777777788888999999
| +4 | 000000001111223333444445555555666666777777788888999999
| +5 | 022344555578899
| +6 | 100237
| +7 | 1
| +8 | 1
| +9 | 1

*Note.* The correlation coefficients $r$ ranged from -.30 in the first row of the display to +.67 in the last row.
$p < .0001$), Creative Achievement Type ($QB[8] = 114.381$, $p < .0001$), and Predicted Time Period ($QB[5] = 115.192$, $p < .0001$) were significant, implying that the mean correlation coefficients of the different classes or levels for each moderator differed from each other. Variability in the magnitude of correlation coefficients between IQ and Creative Achievement for Gender ($QB[2] = 127.385$, $p < .0001$), DT Test ($QB[6] = 464.785$, $p < .0001$), Creative Achievement Type ($QB[8] = 274.390$, $p < .0001$), Predicted Time Period ($QB[5] = 362.915$, $p < .0001$), and Creativity Subscale ($QB[6] = 132.310$, $p < .0001$) were significant, implying that the mean correlation coefficients of the different classes or levels for each moderator differed from each other.

**MODERATOR ANALYSES**

**Gender**

As Table 1 shows, the mean correlation coefficients between DT test scores and Creative Achievement were higher than the mean correlation coefficients between IQ and Creative Achievement for all of the three groups (male, female, and combined). The heterogeneities observed within the different levels or classes imply that the mean correlation coefficients for each level or class cannot be adequately described with a single correlation coefficient. In other words, the variation in the relationships is due to one or more additional factors that the levels or classes do not capture. There were no statistically significant differences between males and females for the mean correlation coefficient between IQ and Creative Achievement ($p = .350$), whereas the mean correlation coefficient between DT test scores and Creative Achievement for males was higher than for females ($p = .0005$).

**TABLE 1. Gender as a Moderator.**

<table>
<thead>
<tr>
<th>Gender</th>
<th>IQ &amp; Creative Achievement</th>
<th>DT test scores &amp; Creative Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (# of r)</td>
<td>Mean r</td>
</tr>
<tr>
<td>Male</td>
<td>15</td>
<td>.216</td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>.246</td>
</tr>
<tr>
<td>Combined</td>
<td>.130</td>
<td>.193</td>
</tr>
</tbody>
</table>

**Note.** No statistically significant differences between male and female groups ($p > .001$).

Homogeneity: heterogeneous when $p < .001$.

Model for Gender, $QB (2) = 27.582$ ($p < .0001$) for the correlation coefficients between IQ & Creative Achievement; $QB (2) = 127.385$ ($p < .0001$) for the correlation coefficients between DT test scores & Creative Achievement.
IQ Tests and DT Tests

As Table 2 shows, the contrasts between IQ Tests revealed that the correlation coefficient between IQ Tests and Creative Achievement for Lorge Thorndike Intelligence Tests as a measure of IQ was statistically significantly higher than the correlation coefficients for either California Test of Mental Maturity (\(p < .0001\)) or Terman Concept Mastery Test (\(p < .0001\)). The contrasts between DT Tests revealed that the correlation coefficients between DT Tests and Creative Achievement for TTCT, GIFT, and WAT were statistically significantly higher than the correlation coefficients for either Guilford Divergent Thinking Tasks (\(p < .0001\)) or Wallach & Kogan Divergent Thinking Tasks (\(p < .0001\)).

TABLE 2. IQ Tests and DT Tests as Moderators.

<table>
<thead>
<tr>
<th>IQ/DT test</th>
<th>N (# of (r))</th>
<th>Mean (r)</th>
<th>Homogeneity</th>
<th>Contrast</th>
<th>(p)-value for Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>41</td>
<td>.073</td>
<td>heterogeneous</td>
<td>California/Lorge</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Lorge</td>
<td>17</td>
<td>.310</td>
<td>heterogeneous</td>
<td>Lorge/Terman</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Terman</td>
<td>32</td>
<td>.180</td>
<td>heterogeneous</td>
<td>Terman/California</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Others</td>
<td>4</td>
<td>.032</td>
<td>homogeneous</td>
<td>TTCT/Guilford</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>TTCT</td>
<td>142</td>
<td>.328</td>
<td>heterogeneous</td>
<td>GIFT/Guilford</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Guilford</td>
<td>38</td>
<td>.108</td>
<td>heterogeneous</td>
<td>TTCT/Wallach</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Wallach</td>
<td>46</td>
<td>.137</td>
<td>heterogeneous</td>
<td>GIFT/Wallach</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Sounds</td>
<td>4</td>
<td>.163</td>
<td>homogeneous</td>
<td>Guilford/WAT</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>GIFT</td>
<td>5</td>
<td>.334</td>
<td>homogeneous</td>
<td>Guilford/WAT</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>WAT</td>
<td>6</td>
<td>.306</td>
<td>homogeneous</td>
<td>Guilford/WAT</td>
<td>(p &lt; .0001)</td>
</tr>
<tr>
<td>Others</td>
<td>33</td>
<td>.186</td>
<td>heterogeneous</td>
<td>Guilford/WAT</td>
<td>(p &lt; .0001)</td>
</tr>
</tbody>
</table>

Note. California = California Test of Mental Maturity; Lorge = Lorge Thorndike Intelligence Tests; Terman = Terman Concept Mastery Test; TTCT = Torrance Tests of Creative Thinking; Wallach = Wallach & Kogan Divergent Thinking Tasks; Guilford = Guilford Divergent Thinking Tasks; Sounds = Sounds and Images; GIFT = Group Inventory for Finding Creative Talent; WAT = Word Association Tests.

Homogeneity: heterogeneous when \(p < .001\).

Model for IQ Tests and DT tests, QB (3) = 102.415 (\(p < .0001\)) for the correlation coefficients between IQ & Creative Achievement; QB (6) = 464.785 (\(p < .0001\)) for the correlation coefficients between DT test scores & Creative Achievement.

Creative Achievement

As Table 3 shows, for general creative achievement and quality, the mean correlation coefficients between Creative Achievement and DT test scores was higher than IQ. On the contrary, for quantity, the mean correlation coefficient between
Creative Achievement and IQ was higher than DT test scores. Contrasts between the coefficients for general creative achievement and quality \((p = .017)\) and quantity \((p = .003)\), and between quality and quantity \((p = .050)\) revealed that no statistically significant differences for the mean correlation coefficients between IQ and Creative Achievement. Whereas the mean correlation coefficients for both general creative achievement \((p < .0001)\) and quality \((p < .0001)\) were statistically significantly higher than for quantity between DT test scores and Creative Achievement.

### Table 3. Creative Achievement Type as a Moderator.

<table>
<thead>
<tr>
<th>Type</th>
<th>IQ &amp; Creative Achievement</th>
<th>DT test scores &amp; Creative Achievement.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (# of ( r ))</td>
<td>Mean ( r )</td>
</tr>
<tr>
<td>a. General</td>
<td>12</td>
<td>.235</td>
</tr>
<tr>
<td>b. Quality</td>
<td>9</td>
<td>.316</td>
</tr>
<tr>
<td>c. Quantity</td>
<td>6</td>
<td>.238</td>
</tr>
<tr>
<td>Contrast</td>
<td>No significant differences((p &gt; .001))</td>
<td></td>
</tr>
<tr>
<td>a. Art</td>
<td>20</td>
<td>.056</td>
</tr>
<tr>
<td>b. Music</td>
<td>7</td>
<td>.210</td>
</tr>
<tr>
<td>c. Writing</td>
<td>12</td>
<td>.172</td>
</tr>
<tr>
<td>d. Science</td>
<td>18</td>
<td>.061</td>
</tr>
<tr>
<td>e. Leadership</td>
<td>4</td>
<td>.365</td>
</tr>
<tr>
<td>f. Social Skills</td>
<td>6</td>
<td>.119</td>
</tr>
<tr>
<td>Contrast</td>
<td>ba, bd, ca ((p &lt; .001))</td>
<td>ea, ed, ef ((p &lt; .0001))</td>
</tr>
</tbody>
</table>

**Note.** Homogeneity: heterogeneous when \( p < .001 \).

Model for Creative Achievement Type, \( QB (8) = 114.381 \) \((p < .0001)\) for the correlation coefficients between IQ & Creative Achievement; \( QB (8) = 274.390 \) \((p < .0001)\) for the correlation coefficients between DT test scores & Creative Achievement.

For art, writing, science, and social skills, the mean correlation coefficients between DT test scores and Creative Achievement were higher than the mean correlation coefficients between IQ and Creative Achievement. On the contrary, for music and leadership, the mean correlation coefficients between IQ and Creative Achievement were higher than the mean correlation coefficients between DT test scores and Creative Achievement. Contrasts among art, music, writing, science, leadership, and social skills revealed that the mean correlation...
coefficients between IQ and Creative Achievement for leadership was statistically significantly higher than art \((p < .0001)\), science \((p < .0001)\), and social skills \((p = .0003)\); music was statistically significantly higher than art \((p = .0004)\), and science \((p = .0008)\); and writing was statistically significantly higher than art \((p = .0005)\). The mean correlation coefficients between DT test scores and Creative Achievement for leadership was statistically significantly higher than music \((p = .0002)\), writing \((p = .0001)\), and science \((p < .0001)\); and art was statistically significantly higher than science \((p = .0002)\). However, the sample set for leadership was limited with only four cases between IQ and Creative Achievement Type and only six cases between DT test scores and Creative Achievement Type, which limits the generalization of the result for leadership.

**Predicted Time Period**

As Table 4 shows, in all time periods except 6-10 years and over 21 years, the correlation coefficients between DT test scores and Creative Achievement were higher than the correlation coefficients between IQ and Creative achievement. In the periods of 6-10 years and over 21 years, the comparison between IQ and Creative Achievement had higher correlation coefficients.

<table>
<thead>
<tr>
<th>TABLE 4. Predicted Time Period as a Moderator.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>a. 0 Year</td>
</tr>
<tr>
<td>b. 1-5 Year</td>
</tr>
<tr>
<td>c. 6-10 Year</td>
</tr>
<tr>
<td>d. 11-15 Year</td>
</tr>
<tr>
<td>e. 16-20 Year</td>
</tr>
<tr>
<td>f. 21+ Year</td>
</tr>
<tr>
<td>Contrast</td>
</tr>
</tbody>
</table>

*Note. Homogeneity: heterogeneous when \(p < .001)\.*

Model for Predicted Time Period, \(QB (5) = 115.192 \(p < .0001)\) for the correlation coefficients between IQ & Creative Achievement; \(QB (5) = 362.915 \(p < .0001)\) for the correlation coefficients between DT test scores & Creative Achievement.

Contrasts among the Predicted Time Periods revealed that for the mean correlation coefficients between IQ and Creative Achievement, the time period of 21 year and more was statistically significantly higher than 0 year \((p < .0001)\), 6-10
year ($p = .0006$), and 16-20 year ($p < .0001$); 11-15 year was statistically significantly higher than 0 year ($p < .0001$) and 16-20 year ($p = .0002$); and both 6-10 year ($p < .0001$) and 16-20 year ($p < .0001$) were statistically significantly higher than 0 year. For the mean correlation coefficients between DT test scores and Creative Achievement, the time period of 11-15 year was statistically significantly higher than 0 year ($p < .0001$), 1-5 year ($p < .0001$), 6-10 year ($p < .0001$), 16-20 year ($p < .0001$), and 21 year and more ($p < .0001$); 0 year was statistically significantly higher than 1-5 year ($p < .0001$) and 6-10 year ($p = .0005$); and both 16-20 year ($p = .0005$) and 21 year and more ($p = .0006$) were statistically significantly higher than 1-5 year.

Creativity Subscale

As Table 5 shows, the contrasts between Creativity Subscale revealed that the correlation coefficient between DT test scores and Creative Achievement for inventiveness was statistically significantly higher than fluency ($p < .0001$), flexibility ($p < .0001$), originality ($p < .0001$), and general creativity ($p < .0001$); creative strengths was statistically significantly higher than fluency ($p < .0001$), flexibility ($p < .0001$), originality ($p < .0001$), and general creativity ($p < .0001$); and elaboration was statistically significantly higher than fluency ($p < .0001$), flexibility ($p < .0001$), originality ($p < .0001$), and general creativity ($p < .0001$).

Multiple Regression Analysis

Variables that were statistically significant moderators (based on the results of the contrasts) were entered into a weighted linear multiple regression model to determine their independent effects for explaining variation in the magnitude of correlation coefficients.

<table>
<thead>
<tr>
<th>Creativity Subscale</th>
<th>N (# of $r$)</th>
<th>Mean $r$</th>
<th>Homogeneity</th>
<th>Contrast</th>
<th>$p$-value for Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Fluency</td>
<td>58</td>
<td>.195</td>
<td>heterogeneous</td>
<td>da, ea, fa</td>
<td>$p &lt; .0001$</td>
</tr>
<tr>
<td>b. Flexibility</td>
<td>37</td>
<td>.205</td>
<td>heterogeneous</td>
<td>db, eb, fb</td>
<td>$p &lt; .0001$</td>
</tr>
<tr>
<td>c. Originality</td>
<td>51</td>
<td>.202</td>
<td>heterogeneous</td>
<td>dc, ec, fc</td>
<td>$p &lt; .0001$</td>
</tr>
<tr>
<td>d. Elaboration</td>
<td>15</td>
<td>.300</td>
<td>homogeneous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Inventiveness</td>
<td>6</td>
<td>.401</td>
<td>homogeneous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Strengths</td>
<td>28</td>
<td>.322</td>
<td>homogeneous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. General</td>
<td>79</td>
<td>.214</td>
<td>heterogeneous</td>
<td>dg, eg, fg</td>
<td>$p &lt; .0001$</td>
</tr>
</tbody>
</table>

Note. Homogeneity: heterogeneous when $p < .001$.

Model for Creativity Subscale, $QB (6) = 132.310 (p < .0001)$ for the correlation coefficients between DT test scores & Creative Achievement.
For the correlation coefficients between IQ and Creative Achievement Type, leadership was reported for only four cases, and the correlation coefficients between IQ and it were homogeneous ($p = .315$). For the correlation coefficients between DT test scores and Creative Achievement Type, leadership was reported for only six cases, and the correlation coefficients between DT test scores and it were homogeneous ($p = .004$). Thus, leadership was included under general creative achievement for the input for multiple regression analyses. For the correlation coefficients between DT test scores and Creativity Subscale, inventiveness was reported for only six studies, and the correlation coefficients between DT test scores and it were homogeneous ($p = .035$). Thus, inventiveness was included under general creativity for the input for multiple regression analyses. For the correlation coefficients between DT test scores and DT test, GIFT was reported for only five studies ($p = .105$) and WAT was reported for only six studies ($p = .872$), and the correlation coefficients between DT test scores and both of GIFT and WAT heterogeneous. Thus, both of them were included under “others”.

The results of the direct entry of the three statistically significant moderating variables of IQ Test, Creative Achievement Type, and Predicted Time Period for the correlation coefficients between IQ and Creative Achievement into a weighted multiple linear regression analysis indicated that Creative Achievement Type and Predicted Time Period independently accounted for variation in $z_r$. The regression model yielded $R^2 = .301$, adjusted $R^2 = .278$, $F(93) = 12.923$, $p < .0001$.

The results of the direct entry of the four statistically significant moderating variables of DT Test, Creative Achievement Type, Creativity Subscale, and Predicted Time Period for the correlation coefficients between DT test scores and Creative Achievement into a weighted multiple linear regression analysis indicated that DT test and Creative Achievement Type independently accounted for variation in $z_r$. The regression model yielded $R^2 = .237$, adjusted $R^2 = .226$, $F(273) = 20.883$, $p < .0001$.

According to Johnson (1993), the tests of meta-analytic predictor’s significance are inappropriate because the error degrees of freedom in the weighted regression procedure of SPSS are too large. Thus, the appropriate corrections in its regression model were made using DSTAT. The results of the weighted multiple linear regression of effect size $z_r$ on moderator variables are presented in Tables 6 and 7.

Table 6 shows the results of the multiple linear regression of effect size $z_r$ on moderator variables for the correlation coefficients between IQ and Creative Achievement (Weighted by Sample Size). Creative Achievement Type ($\beta = -.026$, $z = -4.314$, $p < .0001$) and Predicted Time Period ($\beta = .041$, $z = 5.291$, $p < .0001$) had significant effects on the magnitude of the correlation coefficients between IQ and Creative Achievement. Various IQ tests ($\beta = -.012$, $z = -.820$, $p = .412$) were uncorrelated with the magnitude of the correlation coefficients.

Table 7 shows the results of the multiple linear regression of effect size $z_r$ on moderator variables for the correlation coefficients between DT test scores and Creative Achievement (Weighted by Sample Size). Various DT Test ($\beta = -.051$, $z = -7.898$, $p < .0001$), Creative Achievement Type ($\beta = -.015$, $z = -5.807$, $p < .0001$),
Predicted Time Period ($\beta = -.020, z = -3.097, p < .01$), and Creativity Subscale ($\beta = .010, z = 2.581, p < .01$) had significant effects on the magnitude of the correlation coefficients between DT test scores and Creative Achievement.

**TABLE 6.** Multiple Linear Regression of Effect Size $z_r$ on Moderator Variables for the Correlation Coefficients between IQ and Creative Achievement (Weighted by Sample Size).

<table>
<thead>
<tr>
<th>Moderator</th>
<th>Standardized $\beta$</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ Test</td>
<td>-.012</td>
<td>-.820</td>
<td>$p = .412$</td>
</tr>
<tr>
<td>Creative Achievement Type</td>
<td>-.026</td>
<td>-4.314</td>
<td>$p &lt; .0001$</td>
</tr>
<tr>
<td>Predicted Time Period</td>
<td>.041</td>
<td>5.291</td>
<td>$p &lt; .0001$</td>
</tr>
</tbody>
</table>

*Note.* Overall regression effect = 52.279, $df = 3$, $p < .0001$ (two-tailed).

**TABLE 7.** Multiple Linear Regression of Effect Size $z_r$ on Moderator Variables for the Correlation Coefficients between DT Test Scores and Creative Achievement (Weighted by Sample Size).

<table>
<thead>
<tr>
<th>Moderator</th>
<th>Standardized $\beta$</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT Test</td>
<td>-.051</td>
<td>-7.898</td>
<td>$p &lt; .0001$</td>
</tr>
<tr>
<td>Creative Achievement Type</td>
<td>-.015</td>
<td>-5.807</td>
<td>$p &lt; .0001$</td>
</tr>
<tr>
<td>Predicted Time Period</td>
<td>-.020</td>
<td>-3.097</td>
<td>$p &lt; .01$</td>
</tr>
<tr>
<td>Creativity Subscale</td>
<td>.010</td>
<td>2.581</td>
<td>$p &lt; .01$</td>
</tr>
</tbody>
</table>

*Note.* Overall regression effect = 200.275, $df = 4$, $p < .0001$ (two-tailed).

**DISCUSSION**

The quantitative syntheses of the literature results indicate that the correlation coefficient between DT test scores and Creative Achievement is higher than the correlation coefficient between IQ and Creative Achievement. Thus, it can be concluded that DT test scores account for more variance in the particular achievement measures in this study than IQ scores do. Because the correlation coefficients were heterogeneous, moderator analyses were conducted in order to examine systematic variations in the data.

The results of moderator analyses for the correlation coefficients between IQ and Creative Achievement indicate that correlation coefficient sizes vary systematically across differing kinds of different IQ Test, Creative Achievement Type, and Predicted Time Period, but do not vary across different Gender groups. The results of moderator analyses for the correlation coefficients between DT test scores and Creative Achievement indicate that correlation coefficient sizes vary...
systematically across differing kinds of DT Test, Creative Achievement Type, Creativity Subscale, and Predicted Time Period, but do not vary across different gender groups.

Moderator Analyses of IQ Test for the correlation coefficients between IQ and Creative Achievement indicate that the Lorge Thorndike Intelligence Test may predict Creative Achievement better than some other IQ tests, especially statistically significantly better than California Test of Mental Maturity and Terman Concept Mastery Test. The results of the contrasts revealed that IQ Test, Creative Achievement Type, and Predicted Time Period explain the differences in the correlation coefficients between IQ and Creative Achievement. However, the variance in the magnitude of the correlation coefficients is not significantly explained by the various IQ Test, but it is significantly explained by Creative Achievement Type and Predicted Time Period according to the results of the weighted multiple linear regression, which determines moderators’ independent effects for explaining variation.

The results of the contrasts revealed that DT Test, Creative Achievement Type, Predicted Time Period, and Creativity Subscale explain the differences in the correlation coefficients between DT Test and Creative Achievement. The variance in the magnitude of the correlation coefficients are statistically significantly explained by DT Test, Creative Achievement Type, Predicted Time Period, and Creativity Subscale according to the results of the weighted multiple linear regression, which determines moderators’ independent effects for explaining variation.

**DT Test**

The mean weighted correlation coefficients between TTCT and Creative Achievement are much higher than both Guilford Divergent Thinking Tasks and Wallach & Kogan Divergent Thinking Tasks. This indicates that the TTCT may predict Creative Achievement better than other DT tests. This finding is consistent with many previous studies such as: Torrance’s conclusion (2002) that TTCT scores are good predictors of creative accomplishments later in life; and Plucker’s conclusion (1999), based on a reanalysis of Torrance’s data, that the best predictor for adult creative achievements was TTCT scores and that the TTCT had a correlation with creative achievement almost three times higher than IQ scores. The TTCT is known for having one of the largest norming samples with valuable longitudinal validations (Davis, 1997) and high predictive validity over a very wide age range (Cropley, 2000, for more information, see Kim, 2006).

**Creative Achievement Type**

The mean weighted correlation coefficients between IQ and music are much higher than both art and science. In contrast, the mean weighted correlation coefficients between DT test scores and art are much higher than science. When analyzing the mean weighted correlation coefficients between Creative Achievement Type and both IQ and DT test scores, music is predicted the best by IQ, and all others, including art, writing, science, and social skills, are predicted best by DT
test scores. Leadership was highly predicted by both, however, the small sample size for leadership for both IQ and DT test scores limits any generalization of this result. This finding is consistent with Runco’s conclusion (1986) that particular areas of performance, including writing and art, are more related to divergent thinking than other areas such as music and science.

**Predicted Time Period**

The mean weighted correlation coefficient between IQ and the Predicted Time Period of 21 years and more is much higher than the mean weighted correlation coefficients for the other time periods. However, there were only seven cases reported for the time period of 21 year and more which limits its generalization. In addition, the periods of 6-10 year and 11-15 year, have higher coefficients than both 1-5 year and 16-20 year.

In contrast, mean weighted correlation coefficient between DT test scores and the Predicted Time Period of 11-15 years is much higher than the mean weighted correlation coefficients for the other time periods. In addition, the periods of 16-20 year and 21 year and more, have higher coefficients than both 1-5 year and 6-10 year. The “10-year rule” (Gardner, 1983; Hayes, 1989; Simon & Chase, 1973) may help explain the finding that creative achievement seems to take several years to manifest itself after taking either an IQ or a DT test. However, limited information was presented on what ages the tests were taken which limits this generalization.

**Creativity Subscale**

The mean weighted correlation coefficients between Creative Achievement and both creative strengths and elaboration, are much higher than fluency, flexibility, originality and general creativity. The high correlation of Creative Achievement to the creativity subscale of creative strengths supports Torrance’s conclusion (1979) that creative strengths are the most important predictor of Creative Achievement among the TTCT subscales. Torrance (E. P. Torrance, personal communication, October 30, 2002) also warned that creative strengths are too important to be excluded from full explanations of TTCT scores.

The high correlation of Creative Achievement to the Creativity Subscale of elaboration may support Kirton’s (1976, 1978, 1989) Adaptor-Innovator (A-I) Theory. Recent studies (Kim, 2006; Kim, Cramond, & Bandalos, 2006) explored the possibility of a two-factor model of creativity based on Kirton’s A-I Theory. These studies concluded that a proposed model of Innovators and Adaptors was a good fit. Innovators, creative types who tended to be wildly creative, should score higher on fluency and originality. Adaptors, creative types that tend to adapt existing constructs, should score higher on elaboration and abstractness of titles. The results that fluency and originality had lower relationships with Creative Achievement and elaboration had a higher relationship with Creative Achievement might indicate that Adaptors are more common. However, further studies are needed in order to support this hypothesis.
In conclusion, DT test scores correlated better to Creative Achievement when compared with IQ, this indicates that DT test scores account for more variance in the achievement measures included in this study than IQ and may predict Creative Achievement better than IQ. The various IQ Tests or Gender did not significantly explain the differences in the correlation coefficients between Creative Achievement and either IQ or DT test scores, whereas different DT Tests, Creative Achievement Types, Predicted Time Periods, and Creativity Subscales were consistent moderator variables. The significant independent moderator effect for DT tests indicates that among the DT tests, the TTCT may predict Creative Achievement better than other DT tests, while the Guilford Divergent Thinking Tasks, Sounds and Images, and the Wallach & Kogan Divergent Thinking Tasks had significantly less predictability. Among the Creative Achievement Types, IQ may predict music better than DT test scores. However, all other Creative Achievement Types, including art, science, writing, and social skills, are predicted the best by DT test scores. Leadership is highly predicted by both IQ and DT tests but this generalization is limited by a small sample size. Among the Predicted Time Period, 11-15 year explains the relationship between DT test scores and Creative Achievement the best, however, higher Creative Achievement tends to be measured in all longer time periods for both IQ and DT tests. Therefore, the advantage of using DT tests over IQ tests is shown by the results in that the scores are somewhat independent of the specifics of any one achievement measure or time.

The findings of the present study in terms of predictive validity of IQ or DT test scores are limited its generalization because there were only 94 correlation coefficients between IQ and Creative Achievement, whereas there were 274 correlation coefficients between DT test scores and Creative Achievement. Thus, further studies that report correlation coefficients between IQ and Creative Achievement are needed to confirm the results of this study. In addition, the results of the meta-analyses are only as good as the measures of Creative Achievement used in the underlying studies and some researchers have questioned at least some of the measures.

REFERENCES

References marked with an asterisk (*) indicate studies that included in the meta-analysis. References marked with double asterisks (**) indicate studies that initially qualified for inclusion, but were excluded because of the inability to calculate an effect size.


Dr. Kyung Hee Kim, Assistant Professor, PO Box 8795, School of Education, Jones Hall, The College of William and Mary, Williamsburg, VA 23187, kkim@wm.edu, 757-221-2345