Children's creative potential: An empirical study of measurement issues

Todd Lubart a,⁎, Chantal Pacteau b, Anne-Yvonne Jacquet c, Xavier Caroff d

a Université Paris Descartes, Laboratoire Adaptations Travail-Individu (LATT), 71 avenue Edouard Vaillant, 92100 Boulogne Billancourt Cedex, France
b Laboratoire BIOEMCO, Université Pierre et Marie Curie Paris VI, 4 place Jussieu - Case 120, 75252 PARIS, France
c Laboratoire Psychologie de la Perception CNRS-UMR 8158, Université Paris Descartes, 45 rue des Saints Péres, 75006 Paris, France
d Université Paris Descartes, Laboratoire Adaptations Travail-Individu (LATT), 71 avenue Edouard Vaillant 92100 Boulogne Billancourt Cedex, France

Article history:
Received 28 August 2009
Received in revised form 12 February 2010
Accepted 13 February 2010

Keywords:
Creativity
Potential
Assessment
Giftedness
School children

Abstract

Relations between measures of creative potential and different scoring methods were examined in 154 French schoolchildren. The Test for Creative Thinking – Drawing Production (TCT-DP), parallel lines task from the Torrance Test for Creative Thinking, and an object-based creative drawing task were used. Factor analysis of TCT-DP subscores showed an originality factor and an appropriateness factor. The relations between these factors, judge’s creativity ratings of the same drawings, and statistical originality scores based on the frequency of elements in the drawings were tested by means of a LISREL model. Moreover, TCT-DP scores correlated positively but weakly with performance in the parallel lines task and the object-based drawing task.

© 2010 Elsevier Inc. All rights reserved.

Creativity can be defined as the capacity to produce novel, original work that fits within contextual constraints (Lubart, 1994). Work refers to all types of ideas and productions. This work must be novel in the sense that it goes beyond a replication or copy of that which exists. In the field of giftedness, there has been growing interest in measuring children's potential for creative thinking (Lubart, Georgsdottir, & Besançon, in press; Lucito, 1963; Naglieri & Kaufman, 2001; Treffinger, 1980). In particular, creative giftedness, defined as a high potential for original thought, has been identified as an alternative to intellectual giftedness (Renzulli, 1986; Sternberg & Lubart, 1993). Freeman (2001), for example, in a 27-year longitudinal study on 169 English gifted children contrasted “academically-awarded” children and “creative” children; in addition to the nature of their accomplishments, these two groups differed on their motivation, their integration at school, their personality traits, and their family environment.

However, there is a longstanding debate concerning how creative potential can be measured in order to assess children's natural ability levels, identify creative giftedness and measure effects of educational training programs (Cropley, 2000; Hunsker & Callahan, 1995; Sternberg & Lubart, 1992). Creative potential has, to date, been most often measured by divergent thinking tests, such as the Torrance Test for Creative Thinking (Torrance, 1974). In divergent thinking tests, children are requested to generate as many responses as possible in a limited time. For example, a graphic stimulus, such as a circle, is presented and the respondent must generate as many drawings as possible using the circle as the key element in each drawing during 10 min. A child may generate 15 quick sketches, with the circle being used as a face, the sun, the Earth, the moon, Saturn, a ball, a coin, a road sign, eyeglasses, a wheel, etc. The output is scored for fluency (the number of ideas), flexibility (the number of different categories of ideas), originality (the statistical rarity of an idea based on norms from a large set of children), and elaboration (the extent to which an idea is elaborated with detail). Both verbal and graphic forms of divergent thinking tasks exist.

A second major tradition for measuring creative potential has been the use of constrained production tasks. In these tasks, the goal is to produce a single, developed “work” taking into account a set of imposed constraints. For example, children may make a drawing based on a specified topic, make a collage using a set of provided materials, or write a story based on a specified title. The Test of Creative Thinking – Drawing Production (TCT-DP) is another example because a page with several graphic primitives placed in certain locations (such as a semi-circle placed in the upper left corner of the page) is proposed and the child must make a drawing that incorporates these elements (Urban & Jellen, 1996). There are a variety of scoring systems to evaluate creativity in these production tasks; creativity can be measured by the consensual assessment technique, developed by Amabile (1982), in which qualified evaluators give ratings of the productions (using a 7-point scale from low to high creativity, for example). In this case, the average rating for a production, over several judges, is the creativity score. For the TCT-DP, a system composed of several subscores which measure specific characteristics relevant to creative potential (such as the number of graphic elements presented, for example).
used, unconventionality of ideas, use of humor) has been developed by Urban (1991), (2005) and Urban and Jellen (1996).

These different measurement traditions with their specific scoring methods have co-existed for years. The contribution of the present research is to study relations between different scoring systems using the same creative production task as well as to examine relations between scoring systems across creativity tasks. In this regard, children completed the TCT-DP task, producing two drawings which were then scored using (a) the regular TCT-DP scoring criteria, (b) an originality score based on statistical frequency of ideas within the sample, and (c) judges’ ratings using the consensual assessment technique. Furthermore, a figural divergent thinking task and a creative production drawing task using a set of objects were completed. All tasks were based on the graphic domain of expression to focus attention on similarities and differences between scoring systems while maintaining constant the general domain of creative activity. Previous research suggests that creative potential may be somewhat domain specific, and using tasks across several domains would introduce further variability (Lubart & Guignard, 2004).

The TCT-DP task was chosen as the central task in this study for several reasons. First, this task, introduced a decade ago, has been widely used in numerous studies in several countries (such as Australia, China, France, Germany, India, Korea, Peru, Poland, Portugal, Slovakia, South Africa, United Kingdom, United States) including research on creative giftedness and education (Blumen-Pardo, 2002; Cropley & Cropley, 2000; Jellen & Urban, 1989; Maker, Jo, & Muammer, 2008; Rudowicz, 2004). Second, this task allows several different scoring systems to be applied and compared. Urban and Jellen (1996) proposed a set of specific scoring criteria, mentioned earlier, which include the novelty of various aspects of the drawing production as well as responsiveness to task constraints and complexity of drawing productions. In terms of the consensual assessment of drawings, research on judges’ evaluations of creativity have often suggested that judges value originality, compliance with task constraints and the quality of productions. Judges can evaluate the creativity of drawings made in the TCT-DP task. Finally, statistical originality scores based on the frequency (frequency) of ideas with respect to a sample can be calculated for the TCT-DP drawings and provide a measure that relates closely to the scoring method used typically to score originality in divergent thinking tasks.

We hypothesized that the subscores in the TCT-DP would show an originality component and a task constraint component. These two components, especially the originality component, should relate positively with judges’ ratings of creativity and statistical originality scores calculated for the same drawings. These relations should be observed consistently across parallel forms of the drawing task (TCT-DP form A and form B). Finally, positive, moderately strong correlations between creativity scores for the TCT-DP task and other measures of graphic creative potential should be observed, suggesting a coherent creative capacity in children (Dollinger, Urban, & James, 2004).

1. Method

1.1. Participants

Participants were 154 French school children enrolled in three Parisian public schools (67 boys, 87 girls). There were 100 children in third grade (mean age = 8 years 11 months, SD = 8 months, 43 boys, 57 girls) and 54 children in fifth grade (mean age = 11 years 11 months, SD = 8 months, 24 boys, 30 girls). Parents’ and children’s informed consents were obtained for all participants.

Three adults (female teachers and researchers, ages 37, 46 and 49) each with more than 10 years of professional experience concerning children’s artistic creative productions, rated the creativity of children’s drawing productions.

1.2. Measurements

1.2.1. Test for Creative Thinking —Drawing Production (TCT-DP)

This task consists of a sheet of paper with six graphic elements (a semi-circle, a dot, a dashed line, a 90-degree angle, a curved line, a small open square) placed at fixed locations on a page. All of the elements except one (the small open square) are enclosed in a large rectangular frame. Participants must produce a drawing using these elements. The two parallel forms (A and B) differ by the initial orientation of the page placed before the participant; form B is a 180-degree rotated version of form A. A pencil and eraser were provided for this task. Children were instructed to create a drawing that would be original, interesting and different from the drawings that others would produce.

This task was scored in several ways. First, Urban and Jellen (1996) proposed a scoring system based on several specific criteria to be evaluated. These are (a) the number of graphic elements used among the initial elements proposed [CN]; (b) the number of graphic elements used in a meaningful way [CM]; (c) the number of new items added to the composition [NE]; (d) the number of contacts established between the initial graphic elements [CI]; (e) the degree to which the elements were connected thematically [CTH]; (f) use of the element outside the frame [BFD]; (g) use of added elements outside the frame [BF]; (h) use of three-dimensional drawing techniques [PE]; (i) creation of a humorous or emotional atmosphere [HU]; and (j) use of unconventional, non-stereotyped content or graphic forms [UCT]. As proposed by Urban and Jellen, each criterion is scored using a specific notation system, using in general an ordinal measurement scale (for scoring details, see Urban & Jellen, 1996). The speed score (time) was not used in this study.

Second, we created a statistical originality score for each drawing by calculating the frequency with which each graphic element was used in a certain way over the total set of drawings, with separate calculations for form A and form B. For example, the semi-circle in form B was often used to make a face, and this frequent use received zero points (attributed for all uses with a frequency greater than 5%), was used more rarely to make a ball (1 point, frequency of use between 5% and 2% in the sample) and was used rarely to make a ladybug (2 points, frequency less than 2% in the sample). The total originality score for each drawing thus varied from 0 to 12 (6 elements at 2 points each). It should be noted that the frequencies were calculated across all children in the study because initial analyses showed that separate scoring systems for third graders and fifth graders were highly similar.

Third, the creativity of TCT-DP drawings was assessed by three independent judges using the consensual assessment technique proposed by Amabile (1982). Judges were instructed to use a 7-point scale to assess creativity with no prior definition of that characteristic. Drawings were evaluated against each other for the entire sample, with the set of drawings form A evaluated distinctly from the set of drawings form B.

1.2.2. Torrance Tests of Creative Thinking (TTCT) figural tasks

The parallel lines task from the French version of the TTCT was used to measure creative divergent thinking (Torrance, 1976). In this task, participants must make as many drawings as possible using a set of parallel lines as the basis for each drawing. The instructions indicate that the child should produce as many different, original ideas as possible, within a 10-minute time limit. Fluency, flexibility and originality scores were calculated using the French norms for elementary school children (Torrance, 1976).

1.2.3. Object-based drawing task

In this task, inspired from Getzels and Csikszentmihalyi (1976) and Lubart (1994), participants are presented with a set of actual objects. They must include at least five of these objects in their drawing production. The objects were: a wooden manikin, a red spatula, a wooden spoon, a cardboard box, a pink stuffed dolphin, a blue light bulb,
and a yellow toy construction hat. They were instructed that they can examine the objects, modify them in their drawing and add other non-provided elements in their composition. A large sheet of paper (A3) and pencils and colored markers were provided. Children were instructed to create a drawing that would be original, interesting and different from the drawings that others would produce.

1.3. Procedure

First, informed consent was obtained from parents and children who accepted to participate in the research. The study was conducted in elementary schools in a resource room that children occasionally used for special activities. Children were tested in small groups (4 to 6 children) with each child working at a separate table. First the TCT-DP form A was completed, followed by form B. Fifteen minutes was allowed for each drawing form. Then participants completed additional tasks. One subgroup of the participants (53 children: 22 third graders, 31 fifth graders) completed the parallel lines task during 10min. Another subgroup completed (45 children: 22 third graders and 23 fifth graders) the object-based drawing task using a set of objects during 20min.

Appropriate scoring procedures were applied to these tasks. As described previously in the materials section, TCT-DP drawings were scored using (a) Urban and Jellen's criteria, (b) a statistical originality score, and (c) using the consensual assessment technique with three judges. An experienced creativity researcher (co-author of this paper) scored the TCT-DP criteria using the established scoring system (To estimate scoring consistency, a second researcher (co-author) assessed a subset (n = 28) of the productions. The interjudge reliability for the TCT-DP criteria showed a median alpha of .92.) In the consensual assessment technique, three judges evaluated independently each drawing (forms A and B) within the context of the total set of productions and ratings are made using a 7-point scale ranging from low (1) to high creativity (7). Within each form (A and B), drawings were presented in a different randomized order for each judge. The judges’ scores for each drawing were then averaged. The parallel lines task from the TTCT was scored by an experienced, trained researcher for fluency, flexibility and originality using the established test scoring norms. Finally, the object-based drawing task was scored for creativity using the consensual assessment technique with three judges, with the drawings presented in a different randomized order for each judge. Judges showed adequate inter-rater agreement for the TCT-DP and object-based creativity ratings (Drawing form A α = .78; Drawing form B α = .75, object-based drawing α = .87).

2. Results

The results will be presented in two sections. The first section focuses on the TCT-DP task, for which all participants completed the two parallel forms, which were scored using Urban and Jellen's system, a statistical originality-based system, and by the consensual assessment technique.

The second section of the results presents evidence for links between different kinds of artistic creativity tasks. In particular, relationships between TCT-DP form A drawings, TCT-DP form B drawings, creative divergent thinking in the parallel lines task, and the integrative-productive object-based drawing task are presented.

2.1. TCT-DP: relations between scoring systems

In a first group of analyses, we examined the relations between the different subscores used in Urban and Jellen's scoring system for the TCT-DP. In total, ten subscores were obtained for each of the two TCT-DP drawings (forms A and B). We conducted factor analyses for the rating scores obtained for each drawing and then for the combined set of subscores from both drawings together. The results of principal components analyses were very consistent for drawing form A and drawing form B. The results of the factor analyses are presented in Table 1. A two-factor solution was obtained using traditional Kaiser and Scree test criteria, with approximately 57% and 53% of the variance explained for drawing forms A and B respectively. The results, after varimax rotation, show that the first factor was centered on the number of graphic elements used, meaningful use of these elements and use of the outside element. These scores all relate to providing an appropriate response with respect to the task instructions to use the provided elements. The second factor concerned the use of new elements, contacts and thematic connections between the elements, unconventional or non-stereotyped use of the graphic elements, presence of humor and ambiance, and use of three-dimensional techniques. The scores involved in this second factor relate to the production of a coherent, original and developed artistic idea that underlies a drawing production. Thus, the connection between elements, relates to a creative synthesis, new elements and unconventional use of the provided graphic elements concerns the novelty of this synthesis (with humor as an example of a novel conceptual syntheses and three-dimensional elements in terms of technical quality). This two-factor structure, found for both drawing form A and drawing form B, corresponds to the main components of most creativity definitions, namely an originality-novelty component (factor 2) and an adaptiveness component that concerns respect of task constraints (factor 1). Using the subscores that showed consistent strong saturations (>.60) in the factor analysis, a composite score was calculated for the originality factor and the adaptiveness factor. These factor-based scores were used in subsequent analyses.

In a second group of analyses, we examined the extent to which two other measures of creativity based on the same TCT-DP drawings productions yield similar results. The first alternative score was based on statistical originality of the use of each graphic element in the final drawing. Frequencies for the total sample were calculated for drawing form A and for drawing form B. The second alternative score was judges' evaluations of artistic creativity. Judges reported, in general, at the end of the rating task that their evaluations of creativity were based on novelty in drawing content, use of the graphic elements provided in the task, and mastery of artistic techniques. The correlations between these alternative scores and the TCT-DP subscores are presented in Table 2. These correlations were analyzed using LISREL. The model which was tested posited that the different scores measured three latent variables: originality, adaptation and

| Variables | Drawing form A | | | Drawing form B | | |
|-----------|---------------|-----------|-----------|---------------|-----------|
| | Factor 1 | Factor 2 | Expl. Var. | | Factor 1 | Factor 2 | Expl. Var. |
| Cn | .99 | .05 | .98 | .99 | .07 | .94 |
| Cm | .65 | .45 | .63 | .56 | .24 | .74 |
| Ne | −.03 | .62 | .38 | −.08 | .76 | .62 |
| Ci | .30 | .65 | .51 | .13 | .62 | .39 |
| Ch | −.15 | .93 | .88 | .09 | .83 | .61 |
| Bfd | −.77 | −.02 | .50 | .89 | −.02 | .93 |
| Bfl | .47 | .49 | .46 | .33 | .47 | .53 |
| Bh | −.02 | .52 | .27 | −.21 | .52 | .44 |
| Hu | .15 | .76 | .60 | .13 | .69 | .51 |
| Uct | .28 | .55 | .39 | .27 | .49 | .39 |
| Expl. Var. | 2.43 | 3.27 | 2.35 | 2.91 |
| Pct. Var. | 24% | 33% | 24% | 29% |

Cn = Number of graphic elements used among the initial elements proposed; Cm = Number of graphic elements used in a meaningful way; Ne = Number of new items added to the composition; Ci = Number of contacts established between the initial graphic elements; Ch = Degree to which the elements were connected thematically; Bfd = Use of the element outside the frame; Bfl = Use of added elements outside the frame; Bh = Use of three-dimensional drawing techniques; Hu = Creation of a humorous or emotional atmosphere; Uct = Use of unconventional, non-stereotyped content or graphic forms. Expl. Var. = Explained variance; Pct. Var. = percentage of explained variance.
creativity. The statistical originality and composite scores represented the latent «originality» dimension which was calculated for each drawing. In addition, it was postulated that the latent variables of originality and adaptation determined the level of creativity. Fig. 1 presents the path diagram corresponding to this model. The adjustment indices indicate that the model is acceptable ($\chi^2 (17)=43.16, p<.001; \text{GFI}=.94; \text{AGFI}=.87; \text{NFI}=.83; \text{CFI}=.83$). However, the estimated parameters indicate that the three latent variables are more distinct for drawing form B compared to drawing form A. Additionally, the latent variable of originality was better estimated using statistical originality compared to the originality factor score based on several TCT-DP subscores. Concerning the structural model, it was observed that creativity was mainly based on originality rather than involving adaptation, with the latent variables of originality and adaptation themselves being correlated. Of course the nature of the TCT-DP task, which incites participants to provide a drawing that, at least, uses the minimally acceptable number of elements may be important in the interpretation of the relative contribution of the latent variables (originality and adaptiveness) in the current results.

### 2.2. Relations between different kinds of figural creativity tasks

In addition to the TCT-DP task, graphic creative potential was assessed by a divergent thinking task from the TTCT (parallel lines) and the object-based drawing task. Table 3 presents the correlations between measures obtained in these tasks and the TCT-DP. In general, the correlations between the creativity measures across tasks were positive but weak, suggesting that there is a relative specificity in measures of creative potential.

### 3. Discussion

In this research we sought to examine links between different scoring techniques used in the creativity literature. Each technique—judges’ evaluations, statistically-based originality scores, and scores based on specific content-related aspects of a production—tend to be used in an exclusive way in the literature. In other words, most studies use either judges, or statistical originality, or content-related scores. The current study sought to apply and to relate to each other these scoring methods; the TCT-DP drawing task allowed each scoring procedure to be applied. We found that the three scoring methods...
relate moderately well but do not show complete convergence. This is most clear for judges’ scores relating to TCT-DP factor-based scores which focused on novelty and appropriateness; Dollinger, Urban, and James (2004) found in adults similar correlations between TCT-DP scores and judges’ ratings of the same productions. Finally, it is noteworthy that the ten TCT-DP subscores in fact can be relatively well summarized by two factors that reflect originality and adaptiveness in terms of taking task constraints into account. This result contrasts to some extent with previous analyses using data from German children which showed six factors, and data from Chinese children in Hong Kong which showed four factors (Urban & Jellen, 1996; Rudowicz, 2004). However, the subscores loading on Rudowicz’s factors 1 and 4 (Composition and Novelty, Unconventionality and Humor) are united in the originality factor observed in the current study, and the subscores loading on Rudowicz’s factors 2 and 3 (boundary breaking and basic fragment dependent usage) are united in the adaptiveness factor proposed here. Another contribution of the present paper is to show how statistical analysis using structural equation modeling can help to understand how various scores, among the most frequently used in the creativity research domain, are related to creativity by means of its two respective criteria: the originality and adaptiveness of productions. This model, by integrating several measurement methods applied to the same set of productions, thus contributes toward resolving the debate on the nature of creative potential, its definition and components.

A second objective of the current work was to examine how the different scoring methods used with the TCT-DP will relate to scores obtained in other tasks that assess figural creativity. The results show positive but weak correlations, which seem to corroborate the hypothesis of domain specificity of creativity (e.g., Baer, 1998; and for an integrative model, Baer & Kaufman, 2005), at least for productions in the figural domain. Moreover, the various scoring methods used in the present research do not modify this result.

In terms of recommendations, it seems useful for research purposes to evaluate creativity using multiple scoring systems in order to have a profile of scores. For example, judges’ evaluations and statistical originality scores cannot be considered redundant measures. In the same line, to assess creativity, it is useful to use several tasks, such as divergent thinking and constrained production tasks in order to cover as best as possible the domain. If a single task must be used, a production task, which includes both constraints and complexity, evaluated by judges seems most relevant because of its ecological validity. Indeed, the task that most closely models real creative production is the object-based drawing task with judges’ evaluations.

### References


