Assessing Creativity:
The Test for Creative Thinking - Drawing Production (TCT-DP)
The Concept, Application, Evaluation, and International Studies

KLAUS K. URBAN

Abstract

The Test for Creative Thinking - Drawing Production (TCT-DP), its design, concept and evaluation scheme as well as experiences and results of application are described. The test was designed to mirror a more holistic concept of creativity than the mere quantitatively oriented, traditional divergent thinking tests. The specific design using figural fragments is explained. The drawing production is evaluated by means of a set of criteria which at the same time represent the underlying test construct. The test has been normed with various age and ability groups; there were no significant differences between male and female examinees. Various studies with data concerning the reliability and validity of the TCT-DP are provided.

Key words: Identification, Giftedness, Creativity, Test for Creative Thinking - Drawing Production (TCT-DP)
Most traditional creativity tests give mere quantitative information about a very restricted aspect of creativity only. This was specially true for the two creativity tests which were published and available in Germany in the eighties, when first deliberations on the construction of a new instrument started. Both instruments, the "Test zum divergenten Denken [Test for Divergent Thinking]" (TDK4-6; Mainberger, 1977) and the "Verbaler Kreativitätstest [Verbal Creativity Test]" (VKT; Schoppe, 1975), were very much restricted as to their range of applicability. The TDK was standardised only for three grades (4-6), the VKT may be used only with adolescents and adults from age 15 on with good school education, since the test only uses verbal material on a fairly high level. Both instruments are relatively extensive in application and evaluation (time). The TDK lacks a theoretical basis; it consists of not-founded or -commented collection of various creativity tasks from the American tradition, by researchers like Guilford, Torrance, or Wallach and Kogan. The VKT-result is very much dependent on general verbal and intellectual abilities. Furthermore, both tests are speed-tests and thus are close to intelligence tests as far as application is concerned. Both instruments refer solely to divergent thinking and they are restricted only to the aspect of productivity, which, in this case is the quantity of mainly verbally determined ideas.

These limitations in concept and scope of applicability were one reason for Urban and Jellen (1985; 1986; Jellen & Urban, 1986) to develop their new instrument, the "Test zum schöpferischen Denken - Zeichnerisch (TSD-Z)" (Urban & Jellen, 1995), or, in English, "Test for Creative Thinking - Drawing Production (TCT-DP)" (Urban & Jellen, 1996). This assessment device may be seen as an attempt to apply a more holistic and gestalt-oriented approach to diagnostics of creativity. The German term "schöpferisch" was chosen consciously in order to stress the shaping, the production, and the final "gestalt" as the creative end product. We wanted to consider not only divergent, or, still more limited, quantitative aspects, but aspects of quality, like content, "gestalt", composition, and elaboration, too, and other components stressed in the literature, like (mental) risk taking and breaking of boundaries, unconventionality, affection, and humour.

Design and construction

In designing and constructing the new assessment instrument some premises should be noted:

- The test should be applicable to persons of a broad age range.
- It should work as a useful screening instrument in order to help to identify high creative potentials as well as low creative or rather neglected, poorly developed ones.
- The instrument should be simple and economic in application, in conducting, scoring, and interpretation, economic in time and material.
- The test should be highly culture-fair.

2 The TSD-Z is available via www.testzentrale.de or directly from Harcourt Test Services, Frankfurt.
3 The TCT-DP is available from: www.tvtc.com under Children's tests: Thames Valley Test Company: Harcourt Assessment, Halley Court, Jordan Hill, Oxford, OX2 8EJ, United Kingdom, Tel. +44 (0) 1865-888188, Fax -314348, E-mail info@harcourt-uk.com.
Because of a broad applicability, even to young children, and of an optimal culture-fairness we decided to operationalize our concept by means of a drawing production. According to the definition and consistent with the practical applicability of the test, the objectivity of the administration procedure, and for reasons of the availability of materials, of a common perceptual or informational basis, of comparability of the material presuppositions and conditions for different test takers, certain basic stimuli had to be designed and incorporated into the test. One first stimulation for the design came from the additional subtest of the "Heidelberger Intelligenztest (HIT)" (Kratzmeier, 1977). The stimuli mentioned, in the form of figural elements or fragments, were intentionally designed in an incomplete and irregular fashion in order to achieve maximum flexibility as an imperative for creativity. Instead of concepts, symbols, or holistic figures, we decided to use figural fragments which possessed only vague conventional meanings. The completed drawing based more or less (creatively) on these fragments is evaluated by means of a set of categories which represent our theoretical construct of the assessment device.

This important conceptual proposition respectively the "simple", basic, unique design gives rise to a multitude of differing creative responses as seen in our first investigation of gifted and talented student populations as well as in the many thousands of drawings from persons of various age and ability levels from various countries. At the same time, the given figural fragments must have enough suggestibility to trigger more stereotypical responses from students with lower degrees of creativity. These different possibilities for interpretation (i.e., conventional vs. unconventional) lead to a higher "selectivity" and validity of the TCT-DP instrument.

The test asks the examinee to complete a drawing on the basis of some given figural fragments. These six figural fragments of the instrument were designed with the following points in mind: They are (1) different in design, (2) geometric and nongeometric, (3) round and straight, (4) singular and compositional, (5) broken and unbroken, (6) within and outside a (seemingly) given frame, (7) placed irregularly on the space provided, and (8) incomplete.

An additional and extremely important element of the instrument is the "big square frame". Together with the "small open square" outside the large frame this boundary serves the purpose of providing information on the creative component of risk-taking, which was operationalized by us as "Boundary Breaking" in a twofold manner.

The conceptual deliberations led to the following set of fourteen key criteria which constitute as a whole the TCT-DP construct, and also serve as evaluation criteria (Jellen & Urban 1986; Urban & Jellen 1985, 1986):

Continuations (Cn): Any use, continuation or extension of the six given figural fragments.

Completion (Cm): Any additions, completions, complements, supplements made to the used, continued or extended figural fragments.

New elements (Ne): Any new figure, symbol or element.

Connections made with a line (Cl): Between one figural fragment or figure and another.

Connections made to produce a theme (Cth): Any figure contributing to a compositional theme or "gestalt".

Boundary breaking that is fragment dependent (Bfd): Any use, continuation or extension of the "small open square" located outside the square frame.
Boundary breaking that is fragment independent (Bfi) from the "small open square" located outside the square frame.

Perspective (Pe): Any breaking away from two-dimensionality.

Humor and affectivity (Hu): Any drawing which elicits a humorous response, shows affection, emotion, or strong expressive power.

Unconventionality, a (Uc, a): Any manipulation of the material;
Unconventionality, b (Uc, b): Any surrealistic, fictional and/or abstract elements or drawings;
Unconventionality, c (Uc, c): Any usage of symbols or signs;
Unconventionality, d (Uc, d): Unconventional use of given fragments.

Speed (Sp): A breakdown of points, beyond a certain score-limit, according to the time spent on the drawing production.

If creativity transcends chaos through imposition of aesthetic order, then creative process as well as creative product must reflect the character of a "Gestalt" composition or the coherence of an organization. Since "Gestalt" is more than the sum of its parts, the fourteen key criteria cannot stand as single entities. Only as interacting factors do they together reflect a holistic concept of creative thought. In statistical terms, a certain score on a single criterion says nothing about creativity; only the total score for all criteria indicates the value of the creative product.

The subjects are asked to complete the uncompleted drawing somebody else had begun and finished without knowing what would come out of it in whatever way they wish; everything is allowed and correct, they are free to draw how and whatever they wish. The test sheets are collected after completion, the latest after 15 minutes for each drawing. Figure 1 shows two out of 31 drawing examples as published in the manual for illustration.

An extensive description of the evaluation procedure is prepared in the manual; it assists in the assessment of the drawings by means of the above mentioned evaluation criteria and guarantees an objective scoring as far as possible.

Norms for German students from age 4 to 16 or from kindergarten and grade one to grade 10 have been established by means of a large norm sample (N = 2,500), differentiated as to age, grade, type of school. The findings show that the TCT-DP up to the age of 11/12, i.e. grade 5/6, is a developmental test; after that age a kind of plateau begins.

In the meantime norms for pupils from grade 4 to 10 are available for Poland; first norms are established in Korea in the Korean version; Australian norms are in progress, more than 2,000 students from grade 1 to 10 have already been tested there with both test forms.

Empirical studies with the TCT-DP

After several smaller pre-studies the first investigation with the TCT-DP was done with four groups of seventh graders from different academic achievement levels (Urban & Jellen, 1986). As expected the results show a low positive relationship between the average level of academic achievements (i.e. form of school) and the test scores. Nevertheless high creative as well as low creative thinkers were represented in each of the four groups; thus low academic achievers do not necessarily have low creative potential, and high academic achievers do not necessarily display high levels of creativity.
Some high scores were strongly surprising, especially for the teachers of the low level students. For the students with the highest TCT-DP scores in each group additional data, including teacher interviews, were collected. These case studies supported very impressively the hypothesis that the TCT-DP really contributes to the reliable identification of high creative potential and unconventional thinking. The most interesting result was that in each of these cases strikingly non-conforming behavior could be found. In two instances the TCT-DP also contributed to the identification of two subjects who were not only misjudged by many of their teachers but had also been misplaced by ignoring previously obtained testing results.

Since each of the four cases displayed some sort of socio-emotional maladjustment, additional studies with the TCT-DP might substantiate the notion that high creativity may be accompanied by behavioral eccentricities contradictory to social norms and modes. We believe that this so-called "maladjustment" is mainly caused by the unwillingness and inability of the social environment to react favorably and supportively toward seemingly "deviant" but highly creative people.

By means of these high TCT-scores and the case studies, the teachers saw their students with other eyes and no longer considered their deviant behavior as totally negative. This function and effect marks an important field for an educationally oriented application of the TCT-DP.

A comparison of TCT-DP scores with IQ scores within the highest achievement group revealed a correlation of zero. This result, which has to be interpreted cautiously because of a
number of special conditions (such as relative homogeneity of IQ scores and special mathematical talent), adds to the validity of the test in the sense that the TCT-DP measures something different from academic achievement and conventionally understood intelligence. This can provide important clues to the identification of creative and unconventional thinkers among gifted student populations as well as among non-gifted or even underachieving students. The TCT-DP might, indeed, contribute to the culture fair identification of potential knowledge producers, since knowledge production is not only an intelligent act, but also an act of risk taking, unconventionality, and imagination (Jellen & Verduin, 1986). Every one of these three important concepts is embedded in the TCT-DP construct.

The investigations for establishing norms for German population were done between 1988 and 1993. The manual gives an extensive description; some results will be reported shortly. The total distribution of gender in the norming sample (N > 2000) is relatively balanced; 51.05% were male, 48.95% female examinees. There is no significant difference of the means of male and female examinees in the total population. A minimal trend in favour of boys is relatively consistent throughout the various sub-samples, but statistically negligible. Therefore no separate norms were calculated for boys and girls.

High effects could be observed as far as the variable age is concerned. Cautiously the TCT-DP may be labelled a developmental instrument, though the improvements are not linear, but show a step-like course. Interestingly enough, from the age of 11 on no significant changes could be found in the normal school population.

Though single students even from special schools for learning disabled achieve very high scores, in general students from school branches of higher scholastic achievement levels score higher in the TCT-DP. Therefore different norm tables were calculated. A large sample of adults (N>300) was considered to be one population, though the age difference between the students in education and educators/teachers was strong and significant, but their mean scores, which lay slightly higher than that of the Gymnasium students of grade 7, did not differ significantly.

Within the framework of various smaller and larger studies the reliability of the instrument has been examined, too. In various studies the reliability of scoring, the interrater reliability, mostly was found to be above r=.87. The parallel test reliability of r = .62 - .70 is less satisfying, but still acceptable for a test of this kind. The TCT-DP shows a very high differential reliability for the differentiation between the 25% highest and lowest achievers in both testforms (Chi-square = 33.54, C corr.) = .92) as shown with a large Hungarian sample (N = 1100).

Bröcher (1989) used the TCT-DP as a pre- and posttest in his study on creativity training with intellectually gifted students. For the control group with no training he found a very high re-test reliability with the correlation of r=.81 after 8-12 weeks with a mean difference of only 1.5 points. Even for the training group, whose mean average increased by nearly 7 points, a coefficient of r=.71 was calculated.

It is difficult to answer the question of validity for our TCT-DP, since there are no other instruments which are directly comparable to it. Therefore information from various sources and other instruments must be considered. In the sense of discriminant validity zero or negative correlations may support the test’s validity. Thus, for example, relationships with intelligence (traditional IQ-scores) should be very low or only slightly positive for a strongly inhomogeneous (as far as intelligence is concerned) sample. Correlations with pure quantitative
creativity measures or with verbally oriented divergent thinking tests should lie a bit higher, but only low positive.

In fact, the examination of the relationship between intelligence test scores (IST 70) and TCT-scores of the group of mathematically highly talented students from the first study, mentioned above, resulted in a zero correlation. That sample was relatively IQ homogeneous (M/IQ = 140; variation 119-159).

Wolanska & Nečka (1990) in their Polish study calculated the following correlation coefficients with the well-known Raven’s matrices, a non-verbal intelligence test (Table 1). These results fit into usual expectations, especially as far as the lacking, non-significant relationship in the gifted (intelligence homogeneous) group is concerned.

In the meantime some studies by other authors have been done using the TCT-DP. They yielded very interesting results which shed positive lights on questions of validity and reliability.

By means of the TCT-DP Scheliga (1988) compared the (test)creativity of one group, consisting of 28 non- or semi-professional musicians (aged 18 to 35) who are composers, too, with that of another group of 42 persons (aged 20 to 40) working in a scientific-technical institute, who do not make, but love to listen to music. Scheliga found highly significant differences of 14 points (TCT-total) in favour of the musicians’ group (32.7 versus 18.8), thus supporting good evidence for the TCT’s validity.

Herrmann (1987) studied the effects of different training styles with two student soccer teams in regard to various personality traits (from achievement motivation to creativity). Over several years the one team had been trained by its coach in an authoritarian, autocratic way (ATM); the other team by its coach in a more open and democratic style (DTM). According to the expectation that a more democratic education allows for more creative development and behavior, the average score in the TCT-Total of this last group (DTM) is very high and significantly better (28.5 points) than that of the ATM (19.5 points). Here especially the category “New Elements” as well as two others, namely “Humour” and “Connections made by a theme” are responsible for this result.

A study by Jellen & Bugingo (1989) tried to assess the creative potential of engineering students participating in an annual pentathlon competition of a large middle-western university (USA). 34 from about 100 students were tested with the TCT-DP on a voluntarily basis; the 10 main winners of the pentathlon did the test again after the competition. All of them achieved far above average scores; the first winner got the highest TCT-DP score with 61 points. Including qualitative evaluations of the drawing products, too, both authors state that on the basis of their results the TCT-DP seem to be appropriate to identify, even among students of engineering sciences, those with very high creatively productive potentials.

Table 1:
Correlations between intelligence (Raven) and the TCT-DP in one Polish study
(Wolanska & Nečka, 1990)

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>600</td>
<td>.44 ***</td>
</tr>
<tr>
<td>7 - 10 years old</td>
<td>190</td>
<td>.29 ***</td>
</tr>
<tr>
<td>11 - 18 years old</td>
<td>410</td>
<td>.21 ***</td>
</tr>
<tr>
<td>intellectually gifted</td>
<td>108</td>
<td>.14 n.s.</td>
</tr>
</tbody>
</table>
Bröcher (1989) applied the TCT-DP together with the VKT (Schoppe, 1977), a verbal creativity test, and the "Wiener-Matrizen-Test" (WMT), a non-verbal IQ-Test, to a group of intellectually gifted adolescents (average IQ 130), who got mostly verbally oriented creativity training during their participation in a summer course for gifted students. While the VKT correlated significantly with the intelligence test, the TCT-DP did not; it showed a low correlation with the VKT.

In a more recent study factorial validity and relations to consensual assessments supported convergent and discriminant validity of the TCT-DP (Dollinger, Urban & James, 2004). In investigating the psychometric profile of the TCT-DP in the Hong Kong Chinese context, the results of various analyses within the frame of a large study with 2368 students by Rudowicz & Chau (2000) “…provided encouraging empirical evidence for the reliability and structural and discriminant validity of the TCT-DP” (Rudowicz, 2004, p. 202).

In using the TCT-DP Urban (1991) found interesting developmental curves for children from 4 to 8 years of age with a typical "breakdown" for the 6 years old children in school. He formulated the following 6 stages of creative growth which are close to general cognitive development:

1. Autonomous scribbling/drawing: The child is scribbling/drawing something independent from the fragments. He/she is not interested/able (yet) to recognize or perceive the provided information or adapt to the information according to the given problem.

2. Imitation: The child copies the fragments. Beginning accommodation; the child is using the fragments, but without completing or forming or changing.

3. Concluding/Completing: The fragments are completed and become more or less closed, completed, simple figures: circle, square etc. First assimilating but still not very creative drawings.

4. Isolated animation/objectivation: Using own, individual, more complex schemes and assimilation/incorporation of the given fragments by means of creating/interpreting figures as single/isolated objects and/or creatures.

5. Producing thematical relations: Figures, objects are drawn/seen/interpreted as having an inner relation or thematic dependency structure; an intention of shaping/forming/composition becomes recognizable.

6. Formed holistic composition: High stage of creative achievement; all completed/equipped and new elements/parts of the drawing contribute to a holistic composition, to a common meaning/theme, which is expressed, too, by the holistic way of formal figural quality of the drawing (that does not necessarily mean high technical artificial drawing skills).

Böttcher (1993) used the TCT-DP and the recently developed literal-verbal version, the TCT-LV, in a study of environmental conditions of linguistically talented children. Paszkowska-Rogacz (1992) compared creative abilities of deaf students with those of normal hearing pupils; Wołanska & Nečka (1990) from Poland, also investigated the factorial structure and other psychometric features of the test leading to the proposal to norm and to use the TCT-DP as an official screening instrument in Poland for identifying gifted and creative students. Just recently Albrecht (2002) found the TSD-Z (TCT-DP) to be a helpful instrument for professional and career counselling purposes.

Seen as a whole, for all authors and the studies that used the TCT-DP, the feedback is very positive in general and data supporting reliability and validity are mostly very satisfying and encouraging.

Davis (1995) summarizes in his review that the "TCT-DP is a new, carefully developed, and possibly useful creativity test. Some researchers are impressed with its potential for identifying creatively gifted children. It is difficult to assess whether it will do more than existing American creativity tests. Efficiency of administration and scoring is a definite plus" (p.91). And Cropley's (1996) summary of his review reads like this: "The TCT-DP is a major addition to the battery of creativity tests. It offers an approach to creativity tests that goes beyond the divergent-convergent thinking distinction. It also goes some way towards incorporating noncognitive aspects into measurement of creativity. The procedure itself is interesting for the people being tested as well as for those scoring the test. The manual reviewed here is highly readable, and is also thorough, providing not only practical instructions but also convincing theoretical and technical material justifying use of the test by both researchers and practitioners" (p.227).

The test author totally agrees with Davis' final statement: "As with limitations that plague all creativity tests, the TCT-DP should be used in conjunction with other information (e.g., another creativity test, or teachers' or parents' ratings) in order to minimize false negatives - missing creative children whose variety of creativity is different than that measured by a single test" (p.91).

References

