

THE TCT/DP IN HUNGARY : GENERAL FINDINGS

Paper presented at the *Visual Talent Symposium* of ECHA (European Council for High Ability) in Szentendre, Hungary, 26-29 August 1994

by

Viktória Gyebnár and Andrea Kárpáti

Eötvös Loránd University, Budapest

Department of Education

H - 1146 Budapest, Ajtósy Dürer sor 19-21 Hungary

Abstract

This paper discusses the general findings of the Test for Creative Thinking/Drawing Production (TCT/DP) that was administered between 1992 - 1994 to 1050 children and young people aged 7 - 18 as the first phase of the national standardization of this instrument.

In the *Introduction* we describe the content validity of the test as seen from the perspective of art education. Moreover, scoring problems are indicated that question the very high inter-judge reliability ratings reported by the authors. In the *Aims and methods* section, the sample tested is introduced. We also raise the issue of the equivalence of the two parts of the creativity test and suggest that using one version provides the same results so the test may be reduced to one sheet. In the *Discussion*, and results of the individual assessment items are discussed for the different age groups. We compare the developmental model based on results of the German sample offered by the authors of the test. The pattern of development as observed in Hungary is quite different from the German one as it does not show a steady increase of performance. It reveals a decline in creativity at age 12-13 similarly to the Torrance model of creative development while Germans manifest a steady growth after age 8. Moreover, a qualitative analysis of the different subcomponents will be given in comparison to a similar analysis offered by the authors about a group aged 5 - 12. In *conclusion*, results of a small longitudinal study will be given to support the developmental model for the Hungarian sample and underline the importance of longitudinal research for the increase of reliability of the creativity development model provided with the help of TCT/DP.

Introduction to the Test and its Scoring

The Test for Creative Thinking / Drawing Production (TCT/DP) is a creativity test that has two identical parts, A and B, each providing six pictorial fragments (five within a frame, one outside of it) and invites subjects to *complete a picture utilizing the given fragments* within a given time limit. (The Testing Sheets are shown on Figure 1) It is a pen-and-pencil instrument that can be administered individually or in group settings and is standardised for ages 5 - 14 for Germany and the United States. It is intended to serve as a "culture-fair assessment of the creative potential worldwide" and has successfully been introduced already in a variety of different cultures. (¹) We decided to use this test as a measure for creativity in the test battery of the Leonardo Program, a project for the development and assessment of five alternative curricula for the visual arts (²) because it allows the test taker to respond to stimuli in his or her own personal drawing style and build meaningful symbols and images or engage in non-representational design. Most drawing tests induce a certain pattern of visual creation and measure basically fluency. In drawing, it usually means the ability to construct a number of relatively original images and scores are given on the basis of the number of images produced.

¹. The TCT/DP has been used, among other countries, in Cameron, Canada, Germany, Hungary, India, Indonesia, Italy, Nigeria, The People's Republic of China, The Philippines, Poland, Portugal, the Republic of South Africa, Spain, Turkey, United Kingdom the US and Zimbabwe. National standards are published for Germany. The present paper is the first publication on the Hungarian standardisation. On the TCT/DP worldwide cf. : Jellen, Hans G., Urban, Klaus K. (1989)

². On the Leonardo Program cf.: Kárpáti, Andrea (1995)

Fluency, flexibility, originality and elaboration seem to have entirely different meanings for psychologists on the one hand and artists, art teachers on the other hand. The sum total of scores given does not lend itself easily to differentiation in the quality of representation, phantasy and humor. Tests of divergent thinking usually do not favour "visual types": those with an outstanding ability to produce expressive and aesthetic images in a variety of media. The validity of the test from the point of view of art education may be based on the following characteristics:

- the elements are *both culturally and visually neutral* - they do not impose a large number of stereotypical images or emotionally biased readings on the test taker
- *the fragments may be continued in a variety of visual idioms*: they can be included in a realistic picture, in a caricature, in a design or plan, in a decorative ornament or be completed as a non-cohesive set of images or signs. (Cf. Figures 2a -2d)
- *the frame of the test resembles the usual picture frame format* and thus evokes the atmosphere of the normal drawing class and eases the inhibiting "testing situation."
- the assessment items include *homogeneous composition, thematic unity and perspective*
- three qualities that are highly valued by art educators as signs of skill and craftsmanship
- the assessment items include *humour* - a feature that, according to our knowledge, no drawing test ever values - and *creation / utilisation of signs and symbols*. According to art education research, both humour and symbolisation represent a very sophisticated use of visual language.
- *speed* is an inhibiting factor for most visually creative individuals - it is very appropriate that the test gives extra points for speed only if a certain - rather high - level of performance has already been achieved. The *time limit of the test is no constraint*: for most children it is enough to complete the task.

In terms of content validity, the test is remarkable but its scoring instructions may need some further refinement. The authors report very high ratings of inter-judge reliability coefficients (.93 in one and .90 in another study) and discriminant and convergent validities. ((Jellen, H., Urban, K. 1988, Urban, K. 1991). Still, Hungarian psychologists and art education researchers trained in the psychology of vision who scored the tests found several items problematic to judge objectively. We shall return to this problem in the Discussion section later.

Aims and methods

The research reported here is a critical investigation of the utility value and possible modification needs of the Test for Creative Thinking/Drawing Production (TCT/DP) and represents the first phase of the Hungarian standardisation procedure of this instrument. As the test consists of two very similar tasks (completion of fragments on Test Sheet A and B, see Figure 1) we wanted to know if performance in the two versions is significantly different and the utilisation of both tasks was necessary or redundant. Through an analysis of variance of all assessment criteria in all age groups and their correlation with the total score of the test we intended to show the differentiating potentials of the individual criteria and suggest the

omission or modification of those irrelevant for the assessment of creative performance. Furthermore, we also wanted to reveal if and how the performance of the Hungarian children corresponds to German age standards reported by the authors of the test. We revealed which age groups show significant differences and how these differences may be related to the Torrance model of creativity development as observed in Hungary and patterns of drawing performance as reported by standard works on childart. The age groups and school grades of the Hungarian sample are shown on Chart 1.

- please insert Chart 1 about here -

All subjects were tested at the end of the school year (May 1993 , 1994 and 1995) and the scoring was undertaken by two judges - a psychologist and an art educationalist, authors of this paper who consulted a group of experts (educational researchers, art teachers, artists and art students) on problematic issues of the assessment system. 25 of these experts were also trained in the scoring procedure served as third judges for different parts of the sample. After an agreement on disputable items indicated in the Introduction, an average inter-judge reliability coefficient of 88% was achieved.

Results of the test will be discussed in the next section but one finding seems to be directly related to methodological issues of TCT/DP and thus be discussed here. Test for Creative Thinking/Drawing Production consists of two identical tasks with the same sheet used vertically (Test Sheet A) and turned upside down (Test Sheet B) Many of our subjects have found the second task redundant and boring so we compared results of the two subtests through a paired samples t-test in order to find out *if the two subtests actually measured significantly different results that would justify the utilisation of both of them*. We considered items significantly different at $p < 0.05$ and found ***only 3 out of 15 assessment items compared that yielded significantly different results in the two subtests***. These items were the following: Continuation (Cn, $p = 0.002$), Boundary-Breaking Fragment-Independent (Bfi, $p = 0.01$) and Humour (Hu, $p = 0.004$). Bfi is one of the most disputable items in the test and Humour is clearly one of those that may elicit different scores from judges of different temper. As the two subtests yield significantly different results in only 20 % of the assessment items they can be considered identical in terms of creativity measurement. There may still be reasons to administer them both for special purposes, eg. the detection of the relationships of creativity with the Bfd item as the inclusion of the little open square in the whole composition occurs more frequently on Test Sheet B. This sheet elicited more Continuations (Cn) but less expressive, humorous images (Hu). An international comparison of similar analyses is needed to find out *if both or only one, and if so, which of the test sheets should be used in the future*.

Discussion

An important measure of the appropriateness of a test is the *validity of the individual assessment criteria* and the coherence of the test. We calculated the interior correlations of the items on Test Sheet A and B respectively. The results of this analyses are shown on Chart 2a and 2b.

- please insert Charts 2a and 2b about here -

Note: Items ending with B refer to Test Sheet B.

The authors of the test wanted to *produce a measure of creativity that results in a creative work as well*. It seems that they were successful in devising such a task because the most important factors for good results on this test is that the test taker should make use of all fragments through additions and completions, include all of them in a holistic composition and create a strong emotional or humorous effect. The items Cn, Cm and Hu show high correlation values with almost all other items.

We found, however, that there are some items that show very weak relationships or no correlations at all with the others. Cm and Hu do not correlate with Uca. *The most problematic items are, as indicated before, those that are intended to measure the unconventional utilisation of the given fragments*. Important for visual creativity criteria, Ne, Cl and Cth are also unrelated to two of the three Unconventionality items Uca, Ucb and Ucc. . The first sub-item, Unconventional Manipulation (Uca) that rewards the turning and folding of the sheet is closely related to 4 items only (Bfd,Ucb,Ucc and Uctot) and totally independent from the basic items of the test: Nc,Cl,Cth and Hu. It shows no relations with Pe, either. The second sub-item, Abstractness / Fiction (Ucb) will correlate with all other items but three (Ne, Bfd and Pe) , Ucc (Symbol-Figure Combination),is also independent of Pe but the other two items it does not correlate with are very different in nature: Cl and Cth. The last Unconventionality item, Ucd (Non- stereotypical Utilisation of given Fragments /Figures , ironically, will show correlations with all other assessment items but those of its own kind: Uca and Ucb. The three Unconventionality sub-items generate a variable Unconventionality Total (Uctot)- that will already correlate with all other items excluding Perspective We strongly recommend those who use the test not to consider the total figure for Unconventionality only but the parts separately as *the three related items seem to measure different skills and abilities*.

As expected, drawing in perspective (Pe) has very little to do with creativity: it shows no correlations with BfdB, HuB, UcaB, UcbB, UccB and the total score for the Unconventionality items on Sheet B, UctotB. We are inclined to risk the assumption that *perspective drawing - one of the eternal educational goals of European art curricula - seems to be totally unrelated to an important artistic trait: novelty*. We suggest to omit this variable in calculating the total scores of the test though it may be kept as an assessment criterion for those interested in child art development as the correct depiction of space is one of the indicators of the appearance of what is often termed as the "Age of realism" when the school-age child abandons Surrealist modes of depiction. Moreover, the the total figure for the Unconventionality items should not be used as an independent variable as it sums up very heterogenous skills and may falsify correlation results.

The most problematic item - to be discussed later -, Bfd - utilisation of the small, open rectangle outside the frame - has no correlations with Cth,Cl and Pe on the B sheet. Bfi, however, is a better item than our experts presume: it shows strong correlations with all other items but Humour.

The reliability of the scoring guide

Before discussing the test results for the different age groups, let us review the reliability of the scoring guide and scrutinize the *scoring of some assessment items*.⁽³⁾ Abbreviations, numbers and capital letters in brackets indicate the items as abridged and examples as numbered in the test manual.

The mere repetition of a given fragment receives a score under the *Continuation* item and also as a *Completion*. A "very detailed" Completion should be further specified as one including at least six elements. It is often difficult to distinguish between New Elements, Ne and repetitions of Completions thus it should be stated that the latter ones are not considered New Elements at all. This aspect should be added to the manual. Under the item *Connections Made With Lines* (Cl) a score should be given for inclusion. If a large visual element including one or several small elements it is clearly a case of connection by line. (Eg.: fish in an aquarium. It should be included under "*Connections Made That Contribute to a Theme*" (Cth, 5.) that it is not an easy item to score reliably if no title is given. It may be very difficult to interpret the signs and images drawn by children aged 5 - 8 as their symbol systems are often unique and impossible to read on the basis of the knowledge of "adult iconography" - typical meanings associated to certain visual signs. If the item intends to give credit for expressive images, then at least 3 scores should be allotted for individual elements with a powerful atmosphere, even if they are not included in a holistic drawing.

The most disputed assessment items for Hungarian judges are the "Boundary-Breaking items (6. and 7.) You get 6 points for being "*Boundary-Breaking Being Fragment-Dependent*" (Bfd, 6.) if you utilise the open square outside the big frame . Many Hungarian children, however, often interpreted this image as part of the scoring sheet for teachers (i.e. a part of the paper that should not be touched). It would be interesting to see how subjects react if no text was found on the sheets (at present, they include the name of the test and details about the test taker) and thus subjects could perceive the little square clearly as part of the set of images to be completed. Artists and art students, almost without exception decided to ignore it because they considered it disproportionate in contrast to the proportions suggested by the frame. For them, the fragment outside the frame was simply a "*bad Gestalt*" that they ignored not because they were reluctant to break boundaries - their artistic work clearly suggests that they do it all the time - but because its inclusion could easily result in an imbalanced composition. As 3 -6 points may be received for the inclusion of this item, its effect on end results is decisive. We suggest that the weight of scores received for this item on the total test score should be calculated to see its relative value and artists as well as child art researchers should be consulted on the relevance of this item in terms of visual creativity. Our suggestion is to omit it altogether or limit its scores to a maximum of three points. The twin criterion of this, "*Boundary-Breaking Being Fragment-Independent*" (Bfi, 7.) is the item that most artists and art educators find totally unacceptable. It is sometimes hard to distinguish those who disregard the frame because their creativity does not tolerate the boundaries of a given rectangle from those who are simply unable to make a composition within a given frame or slip out because they want to make fun or in apprehension of the constraints of the testing situation.

Perspective (Pe) is generally perceived as an important factor in assessing the developmental level of visual skills but it is disputable if drawing in linear perspective - a *Western European convention* - should be included in a test for assessing creativity. *Humour*

³. The evaluation criteria are described in: Urban (1991), p. 180.

(Hu, 9.) is *extremely culture-specific*: it would be important to have a detailed pictorial scoring guide for this item and a special study to reveal if judges from different cultures are able to value this item with a fair inter-judge reliability coefficient.

Unconventionality (Uc) is a crucially important item for all creativity tests and the assessment guide is very precise with this item. Still, *Unconventional Manipulation* (Uca, 10.a) that rewards the turning of the test sheet is highly disputed among artists who do not consider it an act of visual creativity. It would be interesting to see how this sub-item correlates with the total score of high and low scorers. Our experience is that children who do not understand the task will often utilise the reverse side of the test sheet to do a drawing unrelated to the fragments to be completed and, according to the manual, will be rewarded as unconventional manipulators. Those who fold or tear the sheet do not always strive for a special effect praised by 3 scores in the test- they are simply toying with the paper. It is practically impossible for judges to see the difference between an intentional visual pun and folds done to turn the sheet into a paper airplane.

For the *Abstractness, Fictional* sub-item, (Ucb) depiction of well-known cartoon figures will receive as many scores as signs, symbols and original fictional characters. It is highly disputable if drawing a Mickey Mouse utilising one or more of the fragments is a creative act as creating a new symbol incorporating the fragments given. As for the *Non-Stereotypical Utilisation of Given Fragments / Figures*, (Ucd) a detailed pictorial scoring guide is needed for every country as a part of the national standardisation process of the test because there are significant differences in different cultures and school systems with different art education curricula. Chinese and Hungarian Gipsy children will certainly have different stereotypes in their normal visual language thus they will come up with different from those in the scoring guide stereotypical images.

In conclusion: ***we found assessment items that, according to visual skills research and art /art education theory, should not be considered as measures of creativity*** or at least should not be given the same amount of points as other, much more relevant items. We found other items that All these problems indicate that the training of judges is a crucially important part of the utilisation procedure of the test and, because of the peculiarities of the visual language that constitutes the basis of a drawing test, ***national pictorial scoring guides will be needed to turn this test into a reliable instrument.***

Test results by age groups

We intend to start the discussions of the test results with an analysis of the basic statistics of the age groups for the individual items. This method was called the "qualitative analysis of the test results" by the authors and published for a German sample. (Urban, 1991) Then we shall give the results of the Test for Creative Thinking/Drawing Production (TCT/DP) by age groups and thus show the developmental pattern of the Hungarian sample. that is quite different from the German one as it does not show a steady increase of performance through the ages.

In order to find out how children in our age groups perform in the various different we compared TCT indices by age groups using one-way analysis of variance. ($p < 0.05$) The results are shown on Chart 3.

- please insert Chart 3 about here -

Cn is highest at age 7 and 8 then declines and achieves a high level again at ages 15, 16, 17 - there is no significant difference between the performances of the two age groups. The lowest performance is observed at ages 11 and 13 with a significant increase in performance at age 12 and a similarly significant decline at age 13. Drawing performance as measured by this item is highest at age 18 but the standard deviation is so high that it indicates that only a few 18-year-olds will be able to "climb back the creativity ladder" and achieve the peak of 8-year-olds. This scheme seems to support the U-shaped curve model of children's graphic development described by Gardner and Winner. (Winner, 1982)

Cm is at its highest also at ages 7 and 8, then it declines and will achieve the highest value at age 17. At ages 11 and 13 children produce significantly poorer results as in all other ages. (At age 12, children score higher than at ages 11 and 13 - almost as high as at age 8 - but their performance is still significantly lower than that of all other age groups so we may consider this age also part of the low performance "valley".)

Ne is also highest at ages 7 and 8 and these results are significantly different from those of all the other age groups. There is a big decline in performance at ages 11, 15, 12 and 16 with 11-year-olds producing the worst results. 18-year-olds will score high again - significantly higher than all other age groups except 7-year-olds. Their results are almost as good as those of 8-year-old children but still have significantly lower scores than the 7-year-olds.

For **Cl**, best results were produced by 7- and 8-year-olds - these results will be significantly higher than those of all others except 18-year-olds. The ages of decline is again age 11 with results significantly lower than all others except for those of ages 12 and 13. There is a slow but steady development observable till age 17: there is no significant increase if we compare results year by year but, if we consider 2-3 -year lapses we may observe slight positive changes where development will be mathematically significant. As with all items mentioned before, there is a leap in performance at age 17 and best results are achieved at age 18 - however, the peak of the 7--year-olds will not be reached again.

In **Cth** children will score highest at age 8 and a stagnation stage appears between ages 11 and 17 where children aged 12 receive the lowest score but this score differs significantly only from those of age groups 8, 13, 16 and 18 . Performance is better at ages 18, 16, 13 - in this rank order - but with this assessment item again the level of the 8--year-olds will not be achieved again.

Bfd scores are lowest at age 11 but results are not significantly different from those achieved at ages 7, 8, 13 and 16. Best performance will be given at age 18 - age group 15 approximates it and both age groups have significantly higher performance than the others.

Bfi results achieve peak level at ages 18, 8 and 7. Age 18 shows significantly higher performance than all other ages and represents a developmental jump. The decline area is again between ages 11 and 16 with the lowest performance given at age 11. (With test Sheet B, the increase in performance at age 8 and decrease at age 11 are both stronger.)

REFERENCES

Clark, Gilbert A., Zimmerman, Enid (1992): Issues and Practices Related to Identification of Gifted and Talented Students in the Visual Arts. The National Research Center for Gifted and Talented. *Research - Based Decision Making Series*. The University of Connecticut:Storrs.

Clark, Gilbert A., Zimmerman, Enid (1994): Programming Opportunities for Students Gifted and Talented in the Visual Arts. The National Research Center for Gifted and Talented. *Research - Based Decision Making Series*. The University of Connecticut:Storrs.

Jellen, Hans G., Urban, Klaus K. (1989): Expanding Worldwide Awareness of Creative Potential. *The Gifted Child Today*, Volume Twelve (3),. 31-33.

Jellen, H., Urban, K. K. (1988): Assessing Creative Potential Worldwide: the first cross-cultural application of the Test for Creative Thinking/Drawing Production (TCT/DP). *Creative Child and Adult Quarterly*, 13, 151-167

Kárpáti, Andrea (1995): The Leonardo Program. In: Kauppinen, Herta & Dicket, Mary ed.: *International Trends in Art Education*. NAEA:Washington.

Urban, Klaus K. (1991): On the Development of Creativity in Children. *Creativity Research Journal*, Volume 4 (2) 177-191

Winner, Ellen (1982): *Invented Worlds*. Harvard University Press:Cambridge,etc.

LIST OF FIGURES

1. The TCT/DP Testing Sheets A) and B) . Original size is A/4 for each sheet.
2. Different drawing styles as manifest in the completion of the TCT/DP from the Hungarian sample:
 - a)
 - b)
 - c)
 - d)
3. Stereotypical images of Hungarian children
 - a)
 - b)

LIST OF CHARTS

1. Ages and school grades of the Hungarian sample for the TCT/DP
- 2a, 2b. Interior correlations of the assessment items for Test Sheets A and B
3. Comparing TCT indices by age groups using one-way analysis of variance

Chart 1. *Ages and school grades of the Hungarian sample of the TCT/DP, 1993-94*

GROUP	AGE	GRADE	NO. OF CASES
1.	7	1.	56
2.	8	2.	117
3.	11	5.	75
4.	12	6	176
5.	13	7.	72
6.	15	9.	121
7.	16	10.	153
8.	17	11.	173
9.	18	12.	94

Total number of cases: 1055

Note: The *Hungarian educational system* at the time of taking the tests (spring of 1993 and 1994) was still based on the *8 + 4 model*: an eight-year compulsory elementary school followed by a four-year secondary school. Grades 9.-12. in our sample represent the four grades of secondary grammar school that our sample comes from. In further phases of the standardisation procedure, secondary technical and vocational school students will also be tested to provide a representative sample for age groups 15 - 18.

