Problem Space Matters: The Development of Creativity and Intelligence in Primary School Children

Marisete Maria Welter, Saskia Jaarsveld & Thomas Lachmann

To cite this article: Marisete Maria Welter, Saskia Jaarsveld & Thomas Lachmann (2017) Problem Space Matters: The Development of Creativity and Intelligence in Primary School Children, Creativity Research Journal, 29:2, 125-132, DOI: 10.1080/10400419.2017.1302769

To link to this article: http://dx.doi.org/10.1080/10400419.2017.1302769

Published online: 29 Jun 2017.
Problem Space Matters: The Development of Creativity and Intelligence in Primary School Children

Marisete Maria Welter, Saskia Jaarsveld, and Thomas Lachmann

University of Kaiserslautern

Previous research showed that in primary school, children’s intelligence develops continually, but creativity develops more irregularly. In this study, the development of intelligence, measured traditionally, i.e., operating within well-defined problem spaces (Standard Progressive Matrices) was compared with the development of intelligence operating in ill-defined problem spaces (Creative Reasoning Task) and the development of creativity (Test of Creative Thinking-Drawing Production), in primary school children (N = 175) across four grade levels. Results showed that problem space matters: Traditional intelligence test scores increased regularly with grade level, whereas creativity, as well as intelligence operating in an ill-defined problem space, developed irregularly but similarly. This suggests that when creativity is fostered, intelligence’s ability to operate in ill-defined problem space may be fostered, likewise.

Guilford (1956) defined the concepts of convergent and divergent thinking. He described convergent thinking as the ability to search for a specific solution in well-defined problem spaces, and divergent thinking as the ability to generate ideas of a certain quantity (fluency), variety (flexibility) and exceptionality (originality or relative frequency; cf. Guilford, 1967) in ill-defined problem spaces. Problem spaces are representations of sets of possible and logical steps that one must take to find the final solution. This is true for problems that present themselves as clear and synoptic situations to the one who will solve them, and for problems that were hitherto unperceived and have to be formulated by the perceiver. The former situation requires the ability of problem solving, the latter the ability of problem finding (Carson & Runco, 1999; Kim, 2011; Runco, 1994).

Convergent thinking is considered to be an important subset of intelligence (Sternberg, 1982), whereas divergent thinking is the cognitive process most commonly linked with creativity (Plucker & Renzulli, 1999). Despite the fact that Guilford (1956) considered convergent and divergent thinking as related concepts, intelligence and creativity have been studied as two separate constructs. Here, the focus is on children’s development of intelligence and creativity considering both as two distinct constructs.

Research on the development of intelligence revealed that performances on intelligence tests correlate positively with chronological age and advancement in school education (Brouwers, Van de Vijver, & Van Hemert, 2009; Ceci, 1991; Ceci & Williams, 1997; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Neisser et al., 1996). The latter was found to be the best predictor of performance on intelligence tests (Brouwers et al., 2009). This means that the development of a child’s intelligence is directly affected by school attendance (Ceci, 1991; Ceci & Williams, 1997; Neisser et al., 1996), but also that “schooling fosters the development of cognitive processes that underpin performance on most IQ tests” (Ceci, 1991, p. 703).

Research on the development of creativity revealed less consistent results. Some researchers found a positive correlation between creativity and chronological age or advancement in school education (e.g. Chae, 2003; Claxton, Pannells, & Rhoads, 2005), whereas others even described a decrease in creativity associated with increasing age/grade (e.g. Kim, 2011; Torrance, 1968). A well-known effect of decline in
creative performance with grade level was first highlighted by Torrance (1968). Using a longitudinal design, he found evidence for a decline in creativity in fourth-grade children, which he called the fourth-grade slump. Torrance (1968, 1995) generalized that children in grade 3 are more creative than children in grade 4. According to Nash (1974), this “slump” is a “result of increasing peer group pressure for conforming behavior around age nine or ten” (p. 170). Charles and Runco (2000) found that the fourth grade slump is restricted to certain aspects of creativity; although fluency of ideas increased, the proportion of appropriate ideas declined with grade level.

Several authors emphasized the impact that the educational environment has on creativity (Maker, Jo, & Muammar, 2008), either fostering its development or inciting its decline (Lubart & Georgsdottir, 2004). According to Besançon and Lubart (2008), “The main goals of most educational systems are to transmit knowledge, rigorous working habits, and societal values” (p. 381). These goals form the basis of most educational curricula. Hence, these curricula may place more emphasis on convergent thinking than on divergent thinking, and this may be the reason why the development of creativity and intelligence yield different patterns of development (Houtz, Rosenfield, & Tetenbaum, 1978).

Up to now, however, in studies comparing the development of intelligence and creativity, the intelligence measures considered came from tests in which cognition operates in a well-defined problem space, i.e., where the task is to find the one and only correct answer. Creativity measures, on the contrary, came from tests operating in an ill-defined problem space, where the task is to generate as many ideas as possible. Therefore, it cannot be affirmed whether, if a difference between the development of intelligence and creativity is found, would in fact be due to the application of the different cognitive abilities related to both constructs or rather due to the differences in problem spaces in which these operate. What would be expected when intelligence is measured operating in an ill-defined problem space? Recent research suggests that intelligence can, in fact, be measured and studied operating in ill-defined problem space situations (Jaarsveld et al., 2015).

This research shows that intelligence measured in ill-defined problem spaces functions differently from intelligence measured in well-defined problem spaces (Jaarsveld, Lachmann, Hamel, & van Leeuwen, 2010; Jaarsveld, Lachmann, & van Leeuwen, 2012).

Jaarsveld et al. (2010) inferred that in a creative thinking process, intelligence (considered as mostly convergent thinking) intertwines with creativity (considered as mostly divergent thinking). Others have argued, as well, that creativity works with intelligence (Ambrose, 2009; Cropley, 2006; Runco, 2003). This intertwining between convergent and divergent thinking is called creative reasoning, and it may be understood as the ability to generate original, yet appropriate, solutions (Jaarsveld et al., 2010, 2012).

In other words, both convergent and divergent thinking intertwine within a cognitive process that emerges from a problem task or a problematic situation for which no solution is readily available. Therefore, both abilities mutually contribute to generating a solution, which emerges after several alternating stages of convergent and divergent thinking (Jaarsveld et al., 2015; Jaarsveld & van Leeuwen, 2005). In this process, both abilities are addressed; divergent thinking creates new approaches and ideas, and convergent thinking ensures that the correct choices are taken and that the logical aspects of the solution are considered (Jaarsveld & van Leeuwen, 2005). Aiming to measure this cooperative cognitive process, a new, innovative measurement method, called the creative reasoning task (CRT), was developed (detailed information about the CRT is presented in the method section; see also Jaarsveld et al., 2010, 2012).

In our study, the development of intelligence (operating in closed problem spaces) and creativity in primary school children of grades 1–4 was compared. Both were related to the development of intelligence operating in ill-defined problem space. Taking into consideration findings of research on intelligence and creativity, the hypotheses were: (a) Traditionally measured intelligence will develop continually across grade level and (b) creativity will develop irregularly. The main research question is whether intelligence measured in ill-defined problem spaces, i.e. creative reasoning, develops more like traditional intelligence or more like creativity. In other words, is the difference in development between intelligence and creativity found in previous research due to cognitive abilities related to these constructs or to the problem space in which they operate?

**METHOD**

**Participants**

At the start of the school year, a sample of 175 primary school children from Grade 1 to Grade 4 (between 6 and 10 years old) was tested. State authorities approved this study and parents consented to the participation of their children. Children were informed that they could end their participation at any time during the session if for any reason they felt uncomfortable; one female child did so and, accordingly, her results were excluded from the study. Also excluded were eight children (all Grade 1) who produced a SPM answer form that was invalid because several solutions were marked. Hence, the final sample was composed of 166 children (89 male and 77 female). Per grade, number of participants and mean ages in years were as follows: Grade 1 (M = 6.28, N = 25), Grade 2 (M = 7.00, N = 61), Grade 3 (M = 8.05, N = 39), Grade 4 (M = 9.17, N = 41).
Material

Traditional intelligence and creativity were investigated with the paper-and-pencil version of standardized tests, the Standard Progressive Matrices test (SPM, Raven, 1998) and the Test of Creative Thinking-Drawing Production (TCT-DP, Urban & Jellen, 1995), respectively. Additionally, the CRT (Jaarsveld et al., 2010, 2012) was used to test intelligence operating in an ill-defined problem space.

SPM

The SPM (Raven, 1998) is a nonverbal intelligence test consisting of 60 items grouped in five sets. Each of the items comprises an incomplete pattern presented in the form of a 1 × 1, 2 × 2 or 3 × 3 matrix. Participants are asked to complete the pattern by choosing a figure from six or eight possible solution options given below the matrix. The maximum score is 60 points, because each item is fixed as pass or fail. The items are, at first, easy and simple, but become increasingly more difficult within and across sets, requiring higher levels of cognitive abilities to encode and analyze information (Raven, 2009). The individual test processing time and the difficulty progression of the SPM items are functional in assessing the extent of clear thinking (Heller, Kratzmeier, & Lengfelder, 1998).

The SPM was intended to capture cognitive ability in as many age groups as possible, regardless of education, nationality, or health condition (Heller et al., 1998). Raven intended to develop a test that is easy to administer, is clear to interpret, and could be employed to individuals of different ages (from 6 years on) and socio-economic backgrounds (Raven, 2009), who all have to solve the same set of tasks in the same order (Heller et al., 1998).

TCT-DP

The TCT-DP (Urban & Jellen, 1995) is a test for the measurement of an individual’s creative thinking potential. It can be used to identify very high creative potential, as well as to recognize individuals with underdeveloped creative abilities, who may be in need of stimulation and support (Urban & Jellen, 1995).

The test contains two answering forms (form A and form B) both providing six figure fragments in a square frame, inspiring further drawing. Based on these fragments, the respondent is requested to complete the drawing in a free and open way. The fragments are: semicircle, dot, large right angle, curved line, broken line, and a small open square outside the frame (Urban & Jellen, 1995).

The drawing is evaluated and scored by means of 14 criteria that can deliver up to six raw points each, except for the four criteria of unconventionality which are valued at a maximum of 3 points (maximum score = 72). The criteria are: continuations, completions, new elements, connections made with a line, connections made to produce a theme, boundary breaking that is fragment dependent, boundary breaking that is fragment independent, perspective, humor and affectivity, unconventionality A, unconventionality B, unconventionality C, unconventionality D, and speed. The test is applicable in a single or group testing with individuals aged between 5 and 95 years with administration requiring 15 min.

CRT

The CRT is a diagnostic device that measures intelligence and creativity intertwined in an ill-defined problem space. This is in contrast to traditional intelligence tests that measure intelligence in a well-defined problem space (Jaarsveld et al., 2010, 2012). The CRT involves generating components and generating relationship(s) that connect these components; it thus represents a cognitive thinking process in which both intelligent and creative abilities interact (Jaarsveld et al., 2015). This interaction evolved as a cognitive mechanism when human survival depended on finding effective solutions to both common and novel problem situations (Gabora & Kaufman, 2010; Runco, 2003). Creative reasoning solves that minority of problems that are unforeseen and yet of high adaptability (Jung, 2014).

The CRT is a creative reasoning test because individuals have to reason creatively. They are directed to think creatively because the task comes with an open problem space: The constraints of the task are ill-defined. The instruction asks them “to create a matrix that is as original and as difficult to solve as possible.” Hence, individuals are urged to apply their reasoning abilities in creating a matrix that never existed before but is nevertheless solvable by another person.

The CRT is a creativity test because each individual creates her or his own matrix. The matrix is created on a blank form with only the outlines of a 3 × 3 matrix showing. Each created matrix is different. The frequency distributions of components and relationships created in the CRT do not compare to those of the components and relationships solved in the SPM (Jaarsveld et al., 2010, 2012; Jaarsveld, Lachmann, & van Leeuwen, 2013). In fact, created matrices (Jaarsveld et al., 2015; N = 52, 19–42 years old) were shown features not contained in the SPM, like original intertwining of relationships, components that are not exclusively geometrical, and three-dimensionality.

For this study three test forms were applied, each corresponded to one of the three possible types of matrix formats contained in the SPM: 1 × 1 (continuous pattern); 2 × 2 and 3 × 3. The figure that completes the matrix should be drawn within the matrix in the outlined square in the lower right corner. Children were free in their choice of test form. The CRT contains two subscores: one for intelligence in an ill-defined problem space, i.e., CRT-Relations (CRT-R), and one for creativity, i.e., CRT-Components & Specifications (CRT-C). Due to the research question and design, the latter was not used in the present study.
The CRT-R score represents the logic and coherence in a pattern of components for a matrix that was created and it is evaluated by means of defined relations that can deliver up to 128 raw points. These relations are: matrix $1 \times 1$ (idiosyncratic and semantic coherence, jigsaw, and pattern completion); string (iteration of one component and iteration of two or >2 components); matrix $2 \times 2$ (symmetry, change, increase, and succession); and matrix $3 \times 3$ (change, increase, succession, combination, indication of mathematical operation, and two values, see Jaarsveld et al., 2012 for detailed information). However, as the CRT deals with ill-defined space problems, it is therefore impossible to fix a set of evaluation criteria for all possible solutions.

Procedure
The tests were performed during the first weeks of the school year. Children were tested in groups during class time. All tests were performed within one session. Children were asked to work alone and quietly. To facilitate this, they were seated sufficiently far away from each other. Two researchers conducted the session without the teacher being present. Children were asked first to perform the SPM (45 min), then to generate a SPM-style item in the CRT (20 min), and finally, to complete the TCT-DP (15 min; Form A). The appropriate instruction was given to the whole group before each of the tests. Those children who finished a test before the given time limit were allowed to read books that lay ready to this purpose.

RESULTS AND DISCUSSION
The Kolmogorov-Smirnov test confirmed normal distribution of the data. Therefore, to test the effect of grade level on intelligence (SPM and CRT-R) and creativity (TCT-DP) scores, analyses of variance (ANOVA) were run. Because a Levene’s test revealed that the assumption of homogeneity of variances was violated, the nonparametric Kruskal-Wallis-Test was performed additionally. Moreover, bivariate Pearson’s correlations were performed between grade, gender, and the raw scores of the three dependent variables. Means and standard deviations of the SPM, CRT-R, and TCT-DP raw scores are presented in Table 1.

Results of the Pearson’s correlation showed that grade correlated positively with mean scores of all applied tests, the SPM ($r = .654$, $p < .001$, $N = 166$), the CRT-R ($r = .356$, $p < .001$, $N = 166$), and the TCT-DP ($r = .440$, $p < .001$, $N = 166$). This suggests that children in higher grades performed superiorly in comparison to children from lower grades on the traditional intelligence test, on the intelligence test operating in an ill-defined problem space, and on the creativity test. There was no correlation between the dependent variables and gender.

Partial Pearson’s correlation performed within the whole sample between SPM, TCT-DP, CRT-R raw scores, controlled for grade, revealed no correlation. However, correlation analyses performed separately within each grade level showed significance in Grade 4 ($N = 41$) between the following scores: SPM and CRT-R, $r = .313$, $p$ (two-tailed) $= .046$, $p$ (one-tailed) $= .023$; SPM and TCT-DP, $r = .339$, $p$ (two-tailed) $= .030$, $p$ (one-tailed) $= .155$. For the other grade levels no correlations were found. Similar results were obtained by Jaarsveld et al. (2012; $N = 205$, 6–10 y old; SPM raw scores correlated with those of CRT-R, $r = .192$, $p < .01$ and with those of the TCT-DP, $r = .225$, $p < .001$).

Regarding intelligence, ANOVA results indicated a main effect of Grade on both intelligence test measures: SPM, $F (3, 162) = 43.870$, $p < .001$, $\eta^2_p = .448$, and CRT-R, $F (3, 162) = 10.094$, $p < .001$, $\eta^2_p = .157$. The nonparametric Kruskal-Wallis-Test revealed the same pattern of results as the ANOVA, $H (3) = 79.059$, $p < .001$ for SPM and $H (3) = 28.002$, $p < .001$ for CRT-R.

The SPM mean scores increased regularly through grade levels (see Figure 1a); Bonferroni post-hoc analysis revealed a significant increase from grade 1 to grade 2 ($p = .016$), and from grade 2 to grade 3 ($p < .001$). The increase from grade 3 to 4 reached the 5% criterion of significance for an independent pairwise $t$-test, but it failed after conservative Bonferroni correction. In contrast to the SPM scores, the CRT-R scores improved irregularly with grade level (Figure 1b), i.e., there was a significant increase from grade 2 to grade 3 ($p = .001$) only; no differences were observed between grade 1 and grade 2 and between grade 3 and grade 4.

Regarding creativity, results showed that TCT-DP scores presented a main effect of grade, $F(3, 162) = 13.781$, $p < .001$, $\eta^2_p = .203$. In accordance with the ANOVA, the nonparametric Kruskal-Wallis-Test also showed a significant effect of grade, $H (3) = 35.825$, $p < .001$.

For TCT-DP scores, there was an irregular increase with grade level, similar to the results for the CRT-R (Figure 1c), i.e., Bonferroni post-hoc analysis showed a significant increase only between grades 2 and 3.

Results showed that traditionally measured intelligence develops more regularly across grade levels than creativity. Interestingly, intelligence measured in an ill-defined problem

<table>
<thead>
<tr>
<th>Test</th>
<th>Grade 1 ($N = 25$)</th>
<th>Grade 2 ($N = 61$)</th>
<th>Grade 3 ($N = 39$)</th>
<th>Grade 4 ($N = 41$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPM</td>
<td>24.80 (8.78)</td>
<td>30.48 (9.57)</td>
<td>40.67 (6.43)</td>
<td>43.54 (5.07)</td>
</tr>
<tr>
<td>CRT-R</td>
<td>7.32 (12.49)</td>
<td>6.59 (11.40)</td>
<td>20.77 (18.62)</td>
<td>22.41 (24.63)</td>
</tr>
<tr>
<td>TCT-DP</td>
<td>9.52 (4.74)</td>
<td>11.56 (6.02)</td>
<td>16.59 (6.02)</td>
<td>18.27 (8.80)</td>
</tr>
</tbody>
</table>

Table 1: Means and (standard deviations) of raw scores of SPM, CRT-R and TCT-DP by grade levels.

Note. SPM = Standard Progressive Matrices; CRT-R = Creative Reasoning Test—Relations subscore; TCT-DP = Test for Creative Thinking—Drawing Production.
space showed a pattern of development comparable to the one presented by creativity; a step function where the scores remained the same over the first two grade levels, with a significant peak in grade 3 and no further change in grade 4.

This finding indicates that problem space matters and not necessarily the application of the different cognitive abilities related to the constructs of intelligence and creativity. In other words, it is the character of the problem space, being well- or ill-defined, which offers different possibilities for an individual’s cognitive abilities. In tasks with a well-defined problem space, different abilities are addressed than in tasks with an ill-defined space. Intelligence operating an in ill-defined space seems to compare better to creativity, which also unfolds in ill-defined spaces, than to intelligence operating in a well-defined space. Thus, it seems that cognitive processes addressed and operating in well-defined problem spaces develop differently from those addressed and operating in ill-defined spaces.

**GENERAL DISCUSSION**

Results showed, first, that difference in the development between intelligence and creativity can be ascribed to the differences in problem space in which cognitive abilities operate. Second, when children developed abilities that help them find the one and only correct solution, they do not necessarily develop abilities that help them find new and appropriate solutions.

Traditional intelligence displayed a continual growth throughout grade levels indicating a stable and continuous development. This result is common in the literature; performance on intelligence tests correlates with chronological age and advancement in school education (Brouwers et al., 2009; Ceci, 1991; Ceci & Williams, 1997; McArdle et al., 2002; Neisser et al., 1996; Welte, Jaarsveld, van Leeuwen, & Lachmann, 2016). In a cross-cultural meta-analysis, Brouwers et al. (2009) found that the performance of
children on Raven’s progressive matrices increased across childhood. Findings showed a positive correlation between advancement in school education and test scores; advancement in school education was found to be the best predictor of performance on intelligence tests (Brouwers et al., 2009).

The relationship between advancement in school education and SPM test scores may occur because the SPM requires cognitive abilities that are nurtured and trained in school (Ceci, 1991; Jaarsveld et al., 2012). Alternatively, the relationship could reflect a more general effect of cognitive development: the way the SPM was constructed, i.e., with increasing item complexity, gives benefit to older children who are more advanced in general cognitive development (Jaarsveld et al., 2012). Normally, both factors work together; the older a child gets, the better it can manipulate complex relationships and, the older a child gets, the longer it has benefited from the educational system.

Creativity findings showed only an increase in scores from grade 2 to grade 3. This suggests a possible stagnation in the development of creativity. As reported in creativity research, some decreases in creativity scores may happen in certain stages of the development (e.g., Kim, 2011; Torrance, 1968). Torrance (1968) observed a decline in creativity scores of fourth grade children, the fourth grade slump. Our results cannot be characterized as a slump (decline in score, as found by Torrance), but rather as a slowdown or a stagnation. Charles and Runco (2000) found that the development of creativity in grade 4 depends on what aspects are measured; while the fluency of ideas increases from grade 3 to grade 4, the proportion of appropriate ideas declines. This task involves both aspects, with a focus of fluency (Urban, 2005). This may explain why there is neither a significant increase nor a decrease in the present data.

These results are also in accordance with other studies reporting stagnations in creativity scores for the TCT-DP. For instance, Urban & Jellen (1995) found that, on average, tests scores increase with grade level. However, the improvement was not linear, but showed a step-like pattern with a plateau profile after grade 5. Theurer, Berner, and Lipowsky (2012), tested children in three sessions; at the beginning of the first school year, at the end of the second school year, and at the end of the fourth school year. They found a discontinuous pattern of development of creativity; children’s creative performance stagnated after the end of the second year. Besançon and Lubart (2008) tested 211 children in two sessions. Their results showed an influence of grade on the creative performance. Children who went from third to fourth grade showed less improvement than children who advanced to grade 2 or grade 3.

It can be inferred that traditional intelligence and creativity developed differently across grade levels, with intelligence growing continually and creativity developing in an irregular way presenting peaks and periods of what seemed to be a stagnation. According to Houtz et al. (1978), this discrepancy may be a consequence of educational curricula, which, throughout grade levels, focus increasingly on convergent thinking. Therefore, schools should take reasonable steps to guide children in their creative development (Lubart & Georgsdottir, 2004; Maker et al., 2008), changing from an approach that emphasizes recognition and memorization of information to one that challenges children not only to think creatively and to explore the unknown (Beghetto, 2010; Smith & Smith, 2010; Torrance, 1972, 1987), but also to find and then formulate problems for which no ready available answer is at hand (Carson & Runco, 1999; Getzels, 1987; Kim, 2011; Runco, 1994).

These findings suggest that intelligence measured in an ill-defined problem space developed differently from traditionally measured intelligence, but similarly to creativity. This means that the problem space of tests should be considered in interpreting developmental patterns. The difference found in the developmental patterns of both intelligence measurements confirms what was highlighted by Jaarsveld et al. (2010, 2012): intelligence operates differently in well-defined and ill-defined problem spaces. According to Jaarsveld et al. (2012), this difference may occur because the performance on the SPM requires abilities that are more likely to be trained in school than the abilities required for the CRT in which no solution is readily available. School education mainly aims at developing abilities that help to find the one and only correct solution.

The similarity in development between creativity and intelligence measured in an ill-defined problem space indicates that similar abilities are needed when performing ill-defined tasks; i.e., intelligence operating in an ill-defined problem space requires abilities that typically are requested in creativity tests. This fact may imply that in fostering creativity, the abilities related to intelligence operating in ill-defined problem spaces may be fostered, as well.

This last observation is not only of theoretical relevance, but is also of very practical importance. All observations and implications made about the educational setting in relation to children’s creativity are also valid in relation to intelligence measured in ill-defined problem space situations. In these types of situations, children are encouraged to generate, rather than only solve problems. This was already emphasized by Kim (2011): Children should be encouraged to find and recognize problems, rather than be offered problems to solve.

In sum, these results show that problem space matters: intelligence functions differently in ill-defined and well-defined problem space situations. The cognitive abilities usually required for creativity tasks compare to those that intelligence requires when operating in ill-defined problem
space. The results imply that when education does not attend to creative thinking, it might endanger intelligent thinking which is needed to find and solve problems for which no ready answer is available.

REFERENCES


