

Research Reports

Cross-Sequential Results on Creativity Development in Childhood Within two Different School Systems: Divergent Performances in Luxembourg Versus German Kindergarten and Elementary School Students

Günter Krampen*^a

[a] Department of Psychology, University of Trier, Trier, Germany, and Université du Luxembourg, Luxembourg, Luxembourg.

Abstract

The aim of the study is conducting methodologically sound, cross-sequential analyses of the creativity development of children attending different school systems. Culture-free tests of creativity (ideational fluency and flexibility) and intelligence were administered in 5 cohorts (two kindergarten and first three elementary school years), which were retested in three consecutive years. Samples include 244 Luxembourg and 312 German children enrolled in educational systems with obligatory kindergarten and 6-year comprehensive elementary school *versus* optional kindergarten, 4-year comprehensive elementary school and educational placement thereafter. Results demonstrate (1) linear increases in intelligence, (2) declines of divergent performances after school enrollment in both samples, (3) increases in divergent performances up to the 5th elementary school year in Luxembourg and up to the 3rd elementary school year in Germany (i.e., the next to last school year before educational placement) followed by a second creativity slump. Cross-sequential results confirm discontinuities in the development of divergent productions in childhood.

Keywords: childhood development, creativity, intelligence, school transition, cross-sequential analysis, divergent thinking

Europe's Journal of Psychology, 2012, Vol. 8(3), 423–448, doi:10.5964/ejop.v8i3.468

Received: 2012-05-21. Accepted: 2012-07-10. Published: 2012-08-29.

*Corresponding author at: Universität Trier, 54286 Trier, Germany, email: krampen@uni-trier.de.



This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Empirical results on creativity development are up to now restricted largely to cross-sectional and longitudinal research designs. This is the case for creativity development in childhood as well as in adolescence and adult development with one exception: McCrae, Arenberg, and Costa (1987) administered six measures of divergent thinking to 825 males aged 17 to 101 years and repeated test administration in a subset of 278 males initially aged 33 to 74 after a 6-year interval. While their *cross-sequential results* point at a general decline in divergent thinking for all cohorts tested at a later time, cross-sectional and longitudinal (repeated measures) analyses suggested curvilinear trends, with an increase in scores for men under 40 years and a decline thereafter. Divergent performances refer to open problems with only a few restrictions, for with very many alternative and good solutions are principally possible. Since Guilford (1950, 1967), divergent thinking (i.e., open-mindedness in the arts, inventions, etc.) is contrasted to convergent thinking, which refers to narrow problems with many restrictions and rules, for with the one and only right solution is searched for in problem solving.

Curvilinear developments in divergent thinking are postulated in the 50-year-old hypothesis of [Torrance \(1963\)](#) for childhood as well. Torrance hypothesized creativity development to covary with developmental transitions, for example, the transition from kindergarten education to elementary school education. Torrance developed this hypothesis on the basis of a combination of the results from independent cross-sectional and longitudinal studies on divergent productions of kindergarten and elementary school students. Specifically, he hypothesized that the transition from early childhood (kindergarten) to middle childhood (elementary school) is related to the greater demands on social adjustment and conformity in children's behavior as well as demands on their acceptance of authorities and the requests for convergent thinking, which is frequently thought to be a necessity for the learning of cultural skills like reading, writing, mathematics, etc.

State of Research

In succession, numerous cross-sectional studies (e.g., [Jaarsveld, Lachmann, Hamel, & van Leeuwen, 2010](#); [Klausmeier & Wiersma, 1964](#); [Pickard, 1979](#); [Smith & Carlsson, 1983](#)) and many longitudinal studies (e.g., [Claxton, Pannells, & Rhoads, 2005](#); [Gaspar, 2001](#); [Howieson, 1981](#); [Torrance, 1968, 1981a](#)) found “the creativity slump” in connection with educational transitions, that is, after school enrollment, in the 4th grade. The hypothesis postulated by [Torrance \(1963\)](#) is plausible, empirically supported, and—therefore—its extensive coverage in psychological and educational contexts and publications began very early (e.g., [Getzels & Elkins, 1964](#)), has remained continuous (e.g., [Sternberg & Lubart, 1991](#)) and is current even now (e.g., [Kaufman & Baer, 2006](#); [Kaufman & Beghetto, 2009](#); [Paris et al., 2006](#)). However, three arguments must be considered.

The Necessity of Using Cross-Sequential Studies

Firstly, solid and convincing empirical support for the hypothesis of discontinuities in creativity development in childhood is missing until today—at least, empirical evidence is weak because there are only cross-sectional and longitudinal studies. More methodologically complex and meaningful developmental study designs (than cross-sectional and longitudinal designs) should be implemented, that is, cross-sequential designs permitting the interpretation of cohort and time effects in development within one study. Cross-sectional designs compared age groups at one time of measurement confounding age and birth cohort, longitudinal designs include different times of measurement (in one or more cohorts) confounding age and time effects. The advantage of cross-sequential design is the possibility to identify age effects in direct comparison to time effects and cohort effects (at least, when repeated measurement effects are controlled).

Thus, for childhood creativity development we must empirically test whether the cross-sectional and longitudinal data results on the creativity slump can be replicated and validated by cross-sequential data. For adulthood, [McCrae et al.'s \(1987\)](#) results show that cross-sectional and longitudinal (repeated measures) analyses suggested curvilinear trends (with an increase in scores for men under 40 years and a decline thereafter), findings which, however, do not hold up in their cross-sequential analyses. The cross-sequential results point to a general, rather linear decline in divergent thinking for all cohorts aged initially 33 to 74 years tested six years later.

In addition, cross-sequential analyses should extend the test of the [Torrance \(1963\)](#) hypothesis of a creativity slump following the transition from kindergarten to elementary school education on the transition from elementary school education to secondary school education. This is of particular interest for those national education systems that practice educational placement at the end of elementary school because this soon coming selective educational placement is anticipated by the children, parents, and teachers in the final two years of elementary school education. The anticipation of educational placement is waiting to be managed, it is on the agenda of parents, teachers, and

children. This results in more or less stress, because (convergent) academic achievement and educational degrees are strongly in the focus. In addition, repeated measurement (longitudinal) data should be controlled for retest effects, which is neglected frequently in longitudinal as well as cross-sequential studies.

The Necessity of Using Quasi-Experimental Designs

Secondly, with only a few exceptions, the studies on creativity development in childhood were implemented in only one educational system without any control group. Of course, randomization of children to different educational systems has to be ruled out. However, a quasi-experimental approach comparing childhood cognitive development in different educational systems is possible. There are only a few studies following this research strategy; however, these are restricted to comparisons of school systems and do not refer to different educational administration systems. Early on, [Kogan and Pankove \(1972\)](#) analyzed creativity development longitudinally over a 5-year span in small samples of 29 versus 72 5th graders (final assessment in 10th grade) enrolled in “two separate school systems” ([Kogan & Pankove, 1972](#), p. 428), which are characterized narrowly as a smaller school system A (with individual testing in the sample) and a larger one B (with group testing in the sample). They found “substantial stability in ideational productivity and uniqueness scores” ([Kogan & Pankove, 1972](#), p. 428) over the 5-year period for males in school system A and for females in school system B, however, low developmental stability was found for the females in setting A and males in setting B. Thus the results of [Kogan and Pankove \(1972\)](#) remain unclear.

More recent, [Besançon and Lubart \(2008, p. 381\)](#) presented the results of a “semi-longitudinal study” during two consecutive years with 210 elementary school children with traditional and alternative educational approaches (i.e., *de facto* a cross-sectional study). Their results point to some positive influences of alternative pedagogy on creative development mainly for the Montessori school system, which focuses on discovery teaching methods and self-regulated learning in the tradition of the Italian pediatrician and special education teacher Maria Montessori (1870-1952). [Heise, Böhme, and Körner \(2010\)](#) also compared the development of creativity (as well as intelligence and school achievement) of 98 elementary school children, half of them enrolled either in traditional classrooms or at a Montessori school. Their longitudinal results (4-year time span) reveal a more positive creativity development in the Montessori school children, while the development of intelligence and of school achievement (geometry, arithmetic, and orthography) is very similar in both groups. Thus, there is some quasi-experimental evidence for differences in creativity development in children enrolled in different school systems. However, still missing are quasi-experimental studies comparing childhood cognitive development in different educational systems, which would allow direct interpretations of discontinuities (“creativity slump”) versus continuities in creativity development in childhood dependent on specific transition characteristics of the national educational administration systems under study.

The Necessity of Using Culture-Free, Predominantly Nonverbal Tests

Thirdly, the [Torrance \(1963\)](#) hypothesis refers mainly to tests of ideational fluency, ideational flexibility, etc. as indicators of divergent thinking, which are dominantly verbal not only in test instructions, but in subjects’ answers and their scoring as well. Because of the medium-size correlations between verbal tests of divergent and convergent thinking (both being confounded by language development and social status), they may be interrelated *a priori* in the domain of crystallized intelligence (at least in the low to middle intelligence range; e.g., [Getzels & Jackson, 1962](#); [Hasan & Butcher, 1966](#); [Magnusson & Backteman, 1978](#); [Schubert, 1973](#)), which itself is highly dependent on education and socialization. Therefore, culture-free, predominantly nonverbal tests should be administered. These tests are, however, dominantly verbal in test instructions (and their understanding), but their advantage is that subjects’ test reactions do not require language, because they can be done nonverbally. Such testing allows

a more powerful empirical test while avoiding, to a large extent, confounding biases with reference to language development and social status. Another advantage of culture-free tests is that the correlations between culture-free, predominantly nonverbal intelligence and creativity scales are low and close to zero (e.g., Jaarsveld et al., 2010; Krampen, 1996; Wallach & Kogan, 1965; Wechsler, Oliveira Nunes, Schelini, Ferreira, & Pascoal, 2010).

In addition, results of longitudinal studies point to rather low to medium developmental positional stability of various indicators of creativity in childhood and early adolescence ($.29 < r < .46$; e.g., Gaspar, 2001; Howieson, 1981; Magnusson & Backteman, 1978; Sparfeldt, Wirthwein, & Rost, 2009) in contrast to a higher developmental stability of intelligence (e.g., $r = .75$; Magnusson & Backteman, 1978). Positional stability versus positional plasticity is the *terminus technicus* in developmental psychology for high versus low correlative stability between different times of measurement. The comparable high developmental (positional) plasticity of divergent thinking in childhood and adolescence indicates relevant interindividual differences in the intraindividual developments. To sum up, in light of the empirically weak (supported only by cross-sectional and longitudinal, but not cross-sequential data), yet “often cited 4th grade slump in creativity test scores” (Torrance, 1968, p. 195), the necessity of conducting cross-sequential studies becomes apparent.

The 3-Level Model of Creativity

It should be mentioned that within this primarily descriptive research on creativity development (not only in childhood, but across the whole life span), the longstanding normative approach to creativity is refreshed by the conceptualization of positive psychology (Seligman, 2002; Seligman & Csikszentmihalyi, 2000). Creativity is conceptualized normatively as a one of 24 personal strengths listed in the *Values in Action Classification of Strengths* (VIA-IS; Peterson & Seligman, 2004). Therefore, creativity is in the focus of positive psychology in general (see, e.g., Krampen, Seiger, & Steinebach, 2012; Simonton, 2000) and especially within the context of school education (Kaufman & Beghetto, 2009). This renewal of the normative view on creativity grounded in the research tradition of conceptions, evaluations, and empirical analyses of educational objectives and developmental goals. In this normative context, it must be reflected that all the indicators of divergent thinking and action, which are used mainly in research on childhood development, are the very basics of creativity, but are not identical with creativity: Divergent thinking is a good and necessary, but is not a sufficient prerequisites of creativity. This is explored in a 3-level model of creativity, which was introduced by Treffinger (1980; Treffinger, Isaksen, & Firestien, 1983) and is shown in Figure 1 in its extended modification by Krampen et al. (2012, p. 80).

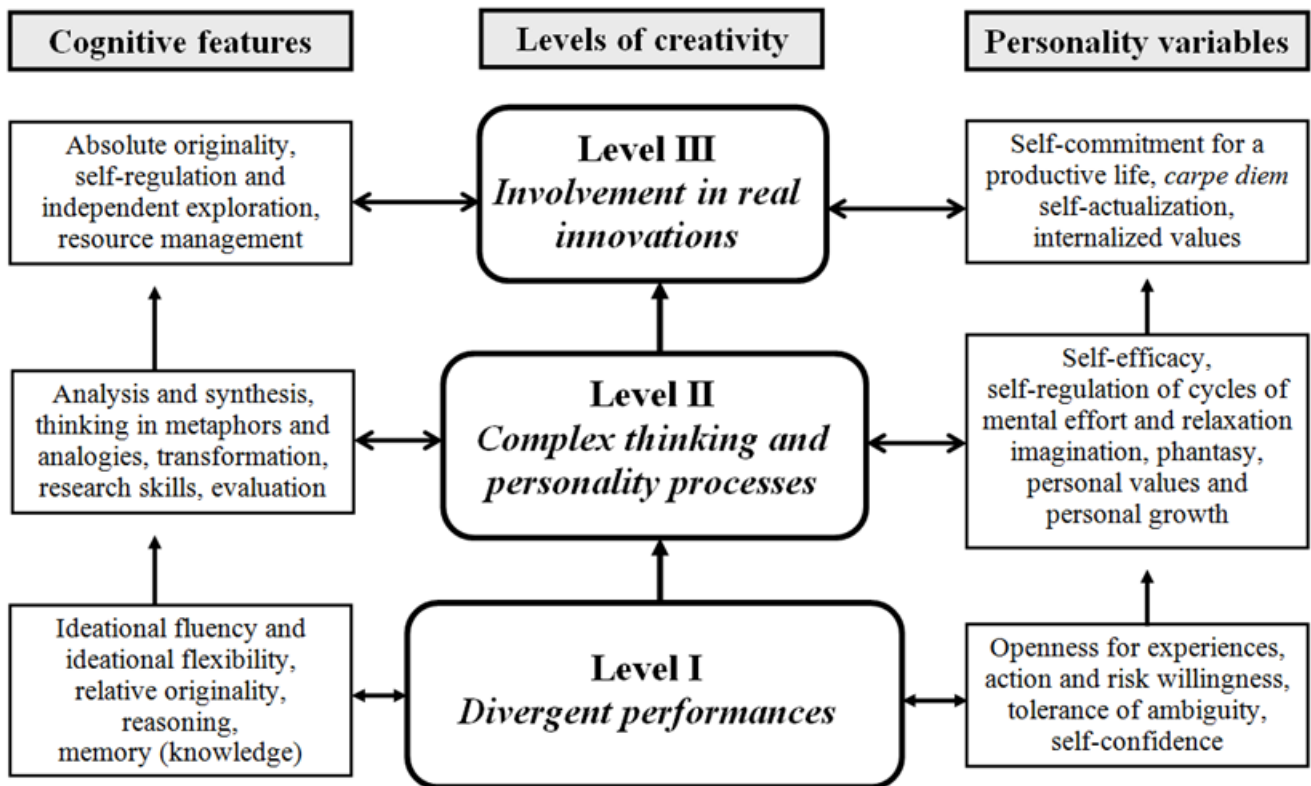


Figure 1. Integrative 3-level model of creativity with attached cognitive and personality characteristics (Krampen et al., 2012, p. 80; extended modification of the proposal by Treffinger, 1980).

Level I of this 3-level model of creativity (Krampen et al., 2012) refers to divergent performances which are related to specific cognitive features, that is, test indicators of ideational fluency and flexibility, relative originality (newness of ideas in comparison to peers), reasoning, and memory, as well as to basic personality variables, that is, openness to new experience, willingness to take risks and action, tolerance of ambiguity, and self-confidence. In developmental research on childhood creativity, it is level I that is the topic of study.

Level II (complex thinking and personality processes) and level III (involvement in real innovations) with their higher-order cognitive features and personality characteristics (see Figure 1) exceed markedly the basic level I. Level II and especially level III refer mainly to the high developmental goal and the personal strength of creativity proposed by positive psychology. However, level I must be evaluated positively as a necessary, albeit not sufficient prerequisite of higher-order, “real” creative productions and innovations. Thus, it is appropriate to analyze the development of divergent thinking skills as early indicators of creativity in children.

The Present Studies: Objectives, Design, and Implementation

Based on the three arguments stated above on the state of research concerning the hypothesis of creativity slumps in childhood development, a cross-sequential study was conducted. Culture-free, predominantly nonverbal psychometric tests on ideational fluency and ideational flexibility as well as intelligence were administered to make powerful empirical tests possible, which avoid the confounding of divergent and convergent test scores with language development and social status of the children. This is especially important because the study was

implemented in two different educational systems in neighboring regions in Western Europe. The first educational system is in Luxembourg. In this system pupils are obligated to attend kindergarten for two years (last two years before elementary school enrollment), a six-year comprehensive elementary school, and educational placement for secondary education after the 6th school year. The second educational system is in Germany, and consists of an optional kindergarten education, a four-year comprehensive elementary school, which is followed by the educational placement into secondary education.

The planning, design, and implementation of the investigations in both countries are identical. However, due to differences in test administration in the two countries, statistical data analyses were conducted separately for each case and results are presented as two independent studies. These differences result from the fact that Luxembourg is a multilingual nation with three official languages (French, German, and Lëtzeburgisch) and multilingual education. Furthermore, approximately 40% of the inhabitants of Luxembourg have a migration background and another primary language (mainly Portuguese, Italian, and English). Therefore, test administration was implemented in the Luxembourg sample by inhabitant and migrant Luxembourg college students in the preferred language of the child. This refers to the verbal test instructions and general communication, but not to children's test answers, because culture-free, predominantly nonverbal tests were used. The German region (which is a direct neighbor of Luxembourg in the open EU countries) is monolingual (German) with only approximately 5% of the inhabitants with a migration background. Therefore, all kindergartens and schools are monolingual (foreign language learning starts at the 5th school year) and all the children were tested in the German language by German college students. Thus, the multilingual test administration in the Luxembourg sample and the monolingual test administration in the German one require separate data analyses including within-sample test score standardizations. Within sample standardizations lead to linear transformed standard values of equivalent Terman-Values (T-values; McCall, 1939). Thus, the identity of the two countries is kept and the comparison of the results in the General Discussion is aimed to the cross-national comparison in which each national context presents some specificity in terms of language and educational system.

The objective of the study reported here was to put the hypothesis of [Torrance \(1963\)](#) to a stringent empirical test, assuring the high standards of developmental psychology research (i.e., cross-sequential design and control of retest effects in repeated measurements), implementing a quasi-experimental approach (which refers to the given reality of two different national educational administration systems in neighboring European countries), and using culture-free, predominantly nonverbal psychometric test indicators of divergent and convergent thinking (for powerful empirical tests avoiding, to a large extent, confounding biases with reference to language development and social status of the children).

Study I: The Luxembourg Study

Method

Sample — In total, 244 kindergarten and elementary school students living in the capital city of Luxembourg (ca. 90.000 inhabitants) as well as surrounding towns and villages participated in the study. They were selected randomly (controlling for sex) from five different educational cohorts: Children were enrolled in the next to the last kindergarten year ($n = 47$), last kindergarten year ($n = 45$) as well as in the 1st ($n = 56$), 2nd ($n = 45$), and 3rd elementary school year ($n = 43$). There was no refusal for participation, neither from children nor from their parents and teachers. Age span at initial test was 4 to 9 years ($M = 6.7$, $SD = 1.57$ years). The sample consists of 124 females (51%) and 120 males (49%).

The children were retested one year as well as two and three years after the initial test (T1). All tests were administered in the winter season after being in the (new) educational cohort, that is, kindergarten year or elementary school grade level, for approximately 10 weeks. Thus, longitudinal data refer to four times of measurement within three years (T1, T2, T3, and T4). This builds up a cross-sequential design with 5 educational cohorts and 4 times of measurement. As a consequence, at T2, children were enrolled in the last kindergarten year up to the 4th elementary school year, at T3 in the 1st up to the 5th school year, and at T4 the data covers the 2nd elementary school year up to the 6th. Age span at T4 was T1 age plus 3 years, that is, 7 to 12 years ($M = 9.6$, $SD = 1.55$). Because of continuous contact to the children and their teachers, dropout rates are small: Retests could be administered at T2 with 235 (96%), at T3 with 232 (95%), and at T4 with 228 children (93%) of the initial sample tested at T1. Dropout analyses did not indicate any biases with reference to the educational cohorts and to the dependent variables tested at the time(s) of measurement before.

In addition, for the control of possible retest effects in test scores, at T2 $N = 115$ children (enrolled in the last kindergarten year and 1st to 4th elementary school year), at T3 $N = 50$ children (1st to 5th elementary school year), and at T4 $N = 88$ children (2nd to 6th elementary school year) were tested for their first time.

Measures — Divergent thinking was tested with a combination of six culture-free, predominantly nonverbal subtests on ideational fluency and flexibility, which were originally developed by Guilford (1964), Mainberger (1977), Torrance (1981b), as well as Acharyulu and Yasodhara (1984) for the assessment of ideational fluency in kindergarten and elementary school students. All subtests were modified for individual test administration without time limit and in scoring for ideational fluency and ideational flexibility by independent test scorers. Subtests are:

(1) *Ovals*: Black ovals on a white page, which should be completed “to do as many different little drawings as possible, which other children will not draw” (at maximum 16 ovals). This instruction invites the child to do something other, more original compared to other children. Ten colored pens are given to the child without time limit for the drawings, which “doesn’t have to be nice, correct, or complete”, but should be denoted by a name/title by the child him/herself in his/her preferred language. Mainberger (1977) constructed this subtest for group testing with time limits for the culture-free, predominantly nonverbal assessment of ideational fluency in elementary school students following the tradition of the well-known “Sketches: Circles” (Guilford, 1964). Subtest *Ovals* was modified for individual test administration with additional picture titling, but without a time limit.

(2) *Picture guessing*: Children were presented with a white page with the beginning of one simple drawing (an uneven, wavy black line in the middle of the page surrounded by a black square). The child is encouraged to guess as many possible completions of this drawing as possible (without a time limit). Mainberger (1977) developed this subtest for group testing with time limit to measure ideational fluency as well. Although it is verbally biased, the child is given the chance to complete the drawing roughly as often as he/she would want to do. Subtest *Picture guessing* was modified for individual test administration, but presented without a time limit as well.

(3) *Free drawings*: The child is instructed to put “as many different little drawings as possible, which other children will not draw” in 24 different fields (squares) on a white page. This instruction invites the child to do something other, more original compared to other children. Again, ten colored pens are given, and it is again emphasized that the drawings “don’t have to be nice, correct, or complete”, but should be denoted by a name/title provided by the child him/herself in his/her preferred language. In the original test by Acharyulu and Yasodhara (1984), these spontaneous drawings are collected from preschool children using clean white pieces of paper (without separated fields) in group testing and with a time limit of 40 minutes. Because our pretests had shown that especially younger

children tend to draw spontaneously only one large picture, the page was structured by marked fields to make the instruction clearer and to encourage children to draw many little pictures. The time limit in individual testing was set to 10 minutes, because pretests indicated that some children (the “painters”) go on drawing for up to 30 minutes or even longer while most of them stopped drawing and were bored within 10 minutes. However, time limit was not communicated explicitly to the children, but (with informal control of time) after a maximum of 10 minutes the child was told that he/she has drawn a lot of pictures and that it is time to change to another game (task).

(4) *Body movements*: This subtest refers asks the child to move “in many different ways” the three-meter distance from point A marked on the floor with red tape to point B marked with blue tape and vice versa. Similar to a subtest of the test battery *Thinking Creatively in Action and Movement* (TCAM; Torrance, 1981b), the child is motivated in individual testing to present many possibilities to bridge the distance by any movement and action. Movements can refer to human ones and to imitations of animals, robots, machines, or work of fiction or fantasy as well. Movements and action should be carried out, but can be verbally described and simulated too. Most children move actively and describe verbally very few and short (e.g., imitations of animal sounds). There is no time limit in individual testing.

(5) *Alternatives actions*: The task is to put plastic cups into a basket “in many different ways, such ways other children will not do” (Torrance, 1981b). Actions must be realized (but may fail); however, they can be additionally described verbally. There is no time limit in individual testing.

(6) *Alternative uses*: The task is to show alternative uses for a beer-mat (beverage coaster). The child gets one and only one beverage coaster each time and is motivated to use it in an alternative, new way. He/she is allowed to do what he/she wants to do with the beverage coaster, for example, snapping, creasing, folding, unpicking, stamping on, smashing, tearing to pieces, etc., but without any aids. Alternative uses of the beverage coaster must be demonstrated (but may fail); however, they can be additionally described verbally. There is no time limit in individual testing. The subtest *Alternative uses* is in the tradition of the well-known “Brick uses/Utility tests” (Guilford, 1964).

Test administration and child’s responses were kept in the test record in full detail. Test scoring was done independently by scorers experienced in the creativity testing of children on the basis of the test record. Test scores refer to ideational fluency and ideational flexibility:

FLU: Ideational fluency is the total number of all ideas presented by a child in all six subtests. All ideas presented in action, movement, and drawings are accepted, independent of verbal comments made by the child. However, simple repetitions of an idea were excluded from scoring as ideas with minimal variations (e.g., subtest *Ovals*: drawings of a cheese pizza, ham and pineapple pizza, salami pizza, etc. are scored together with 1 point only, but Pizza Margarita, Scaloppine Milanese, and Burger with 3 points; e.g., subtest *Body movements*: movements like a dog, like a white dog, like a black dog, etc. are scored with 1 point, but dog, fox, elephant are scored with 3 points).

FLE: Ideational flexibility is the number of different categories which the ideas presented by a child within each subtest belong to. Subtests FLU-scores are added to obtain the total FLU-scale score. The ideas presented in each subtest are classified by the help of inductively constructed categorical systems (Krampen, 1996), which include 7 (*Alternative actions*), 11 (*Alternative uses*), 12 (*Body movements*), or 18 (*Ovals, Picture guessing, Free*

drawings) exclusive and exhaustive categories. For example: Categories for ideas demonstrated in the subtest *Body movements* refer to human movements without aids (e.g., walking, running, creeping, swimming, hopping), human movements with sport aid (e.g., ski, sledges, ice skates), human movements with playthings (e.g., bicycle, scooter, skateboard), movements of animals living on the land (e.g., ape, dog, elephant), movements of birds and other flying animals (e.g., eagle, stork, wasp), movements of animals living in the water (e.g., dolphin, whale, octopus), movements like road vehicles (e.g., automobile, bus, motor scooter, motorcycle), movements like construction plant/machinery (e.g., excavator, crane, bulldozer), movements like rail traffic (e.g., subway, train, streetcar), movements like air traffic (e.g., plane, jet, balloon, rocket), movements like water traffic (e.g., ship, submarine, steamer), and movements in fiction and fantasy (e.g., Superman, Batman, robots, witch, R2D2). The total number of different categories to which the ideas presented by a child belong, is summarized for the six subtests and builds up the FLE-scale-score.

Inter-coder reliability of the FLU- and FLE-scorings was tested by comparisons of independent scorings for the records of the total sample tested at T1 ($N = 244$) and of randomly selected subsamples of $n = 40$ tested each at T2 and T3, respectively. Inter-coder reliability was confirmed by the correlations between two independent scorers for FLU ($r > .97$; $p < .01$) and FLE ($r > .87$; $p < .01$) as well as by the nonsignificant mean differences between the two independent scorings of FLU ($t(242) = 0.74$ and $t(38) < 0.75$; $p > .20$) and of FLE ($t(242) = 0.75$ and $t(38) < 0.69$; $p > .20$).

Reliability of the FLU- and FLE-scores is evaluated by Cronbach's alpha for each of the four times of measurement (see Table 1). The reliability coefficients of FLU and FLE are satisfactory for the intended statistical analyses on the group level (FLU: $\alpha > .83$ and FLE $\alpha > .74$).

Intelligence is measured by two subsets of the *Standard Progressive Matrices* (SPM; Raven, Court, & Raven, 1979), a culture-free, predominantly nonverbal test on inductive reasoning. Because all SPM-tasks require the selection of the one and only correct answer out of wrong alternatives, it is an operationalization of convergent thinking aiming at the single correct solution without any alternatives. In individual testing no time limits were set. In the present study, inter-coder reliability (analyzed by two independent scorers as above), is $r = 1.00$ without any mean differences ($t(xx) = 0.00$). Reliability coefficients for all four times of measurement are very good ranging from $.88 < \alpha < .94$ (see Table 1).

Data on the total *test duration* and on personal characteristics of the child (e.g., age, sex, family size) were also documented in the test record. Test time varied between $M = 64$ ($SD = 6.4$) minutes at T1, $M = 59$ ($SD = 6.5$; T2), $M = 63$ ($SD = 6.2$; T3), and $M = 61$ ($SD = 5.9$; T4) without significant differences between the repeated measurements [$F(3/672) = 2.38$]. There are nonsignificant ANOVA effects of the educational cohort [$F(4/224) = 1.85$] and its interaction with time of measurement [$F(12/672) = 3.37$]. However, there are significant, but rather low correlations of test time with the test scores on ideational fluency (FLU: $.21 < r < .45$; $p < .01$) and ideational flexibility (FLU: $.19 < r < .26$; $p < .01$), but only two significant correlations out of four correlations with the SPM-scores on intelligence (T2 and T4: $.14 < r < .19$, $p < .05$; T1 and T3: $r < .12$). Thus, common variance of the scores on divergent thinking with test duration is at maximum 19%, and the common variance of the scores on convergent thinking with test duration is a maximum of 3%.

Table 1

Coefficients of Internal Consistency (Cronbach's α) at the Four Times of Measurement and Coefficients of Test-Retest Stability Between the Four Times of Measurement in the Luxembourg Sample

Time of measurement	N	Scale		
		Fluency	Flexibility	Intelligence
Internal consistency (α)				
1 st measurement (T1)	244	.84	.76	.91
2 nd measurement (T2)	235	.86	.75	.89
3 rd measurement (T3)	232	.88	.80	.92
4 th measurement (T4)	228	.85	.79	.93
Test-retest-stability (r)				
T1 - T2	235	.49**	.49**	.64**
T1 - T3	232	.32**	.31**	.54**
T1 - T4	228	.30**	.28**	.48**
T2 - T3	231	.48**	.45**	.53**
T2 - T4	228	.29**	.27**	.59**
T3 - T4	228	.27**	.27**	.61**

** $p < .01$.

Results of Study I: Luxembourg

Test-Retest Effects — Checks for retest effects on test scores were analyzed using the data from the $N = 115$ children (enrolled in the last kindergarten year and 1st to 4th elementary school year) tested for their first time at T2, $N = 50$ children (1st to 5th elementary school year) tested for their first time at T3, and $N = 88$ children (2nd to 6th elementary school year) tested at T4 for their first time. Comparisons of these new, independent samples with the longitudinal sample retested at T2, T3, and T4, respectively, did not result in statistical mean differences for neither the test scores on ideational fluency [FLU-T2: $t(348) = 1.15$; T3: $t(280) = 0.85$; T4: $t(314) = 1.03$; $p > .10$] and flexibility [FLE-T2: $t(348) = 0.74$; T3: $t(280) = 0.81$; T4: $t(314) = 0.98$; $p > .10$] nor for intelligence [SPM-T2: $t(348) = 0.55$; T3: $t(280) = 1.01$; T4: $t(314) = 0.73$; $p > .10$] in statistical significant mean differences. Therefore, retest effects in the repeated measurements of divergent and convergent thinking, which may be a result of learning by prior testing or recall of test responses, can be excluded. While children do remember the prior test taking situation, inspection of the contents of random samples of individual test records on FLU and FLE from different times of measurement confirms that children very seldom simply replicate their test reactions (i.e., drawings, movement, actions) presented at prior tests. Rather, children produce new ideas or substantial variations of their prior test responses. A time span of one year between the test administrations seems to be very long for children (for time perception of children see, e.g., Ben Baruch, Bruno, & Horn, 1985; Rattat & Droit-Volet, 2005) and very much happens during one year of a child's life. Therefore, memory for specific test responses and their recall one year later are low.

Positional Stability Versus Plasticity Over Time — Relative (i.e., positional) changes examined longitudinally in the 4-year time span are low to medium in ideational fluency and flexibility, but they are markedly higher in intelligence (see Table 1). Correlation coefficients of intraclass FLU- and FLE-scores, respectively, decline to less than 9% common variance between more distant times of measurement, while common variance of SPM-scores between all different times of measurement exceed 22% and range up to 41%. Thus, high positional stability is confirmed for culture-free tested convergent thinking skills from kindergarten up to the 6th grade level of elementary school education in Luxembourg. On the other hand, positional stability of culture-free tested divergent thinking

skills in childhood development tested one to three years apart is low. This result cannot be explained by general low retest reliabilities of nonverbal creativity tests: For example, Zachopoulou, Makri, and Pollatou (2009) confirmed high retest reliabilities and nonsignificant mean differences of the scales of Torrance's *Thinking Creatively in Action and Movement* (TCAM; Torrance, 1981b) in children for a time interval of two weeks.

Intercorrelations of Divergent and Convergent Thinking Scores at and Between Times of Measurement

— Intercorrelations of ideational fluency and flexibility as well as intelligence at the four times of measurement and between them are presented above the main diagonal in Table 2 for the Luxembourg sample. Cross-sectional correlations (computed with the data from one time of measurement, printed in bold in Table 2) are generally higher than the longitudinal correlations (computed with the data of two different times of measurement). This difference is, on average, statistically significant at $p < .05$. The two nonverbal scales on divergent thinking (FLU and FLE) correlate statistically significantly, with a common variance up to 42%, which confirms earlier results (e.g., Torrance, 1981b; Zachopoulou et al., 2009). The nonverbal test score on convergent thinking (SPM), however, is correlated much lower with FLU and FLE, and the correlations lose, in part, their significance in time-lagged analyses (see Table 2). The trend seems to indicate somewhat slightly higher correlations of convergent thinking (reasoning) with ideational flexibility (FLE) in comparison to ideational fluency, but the common variance of culture-free, predominant nonverbally tested creativity and intelligence in children does not exceed 8%. This is in agreement with results which show that correlations between culture-free, predominantly nonverbal tested intelligence and creativity scales are low and close to zero (e.g., Jaarsveld et al., 2010; Krampen, 1996; Wallach & Kogan, 1965; Wechsler et al., 2010).

Absolute Changes (Mean-Level): Cohort and Time Effects — Cross-sequential data analyses were computed by analyses of variance (ANOVA) with the between-factor *Five Educational Cohorts* and the repeated measurement factor of *Four Times of Measurement*, each being one year apart from the other, thus a 5 x 4 design with repeated measurement of the second factor. Results of the ANOVAs computed for ideational fluency (FLU), ideational flexibility (FLE), and intelligence (IN) are presented for the Luxembourg sample in the upper half of Table 3. Firstly, it should be noted that there are no statistically significant interaction terms. Secondly, all main effects of the between (Cohort) and repeated measurement (Time) factors are statistically significant. In addition, following the terminology of Cohen (1977), for all of these significant ANOVA main effects, large effect sizes (f) are documented, which are even “very large” (almost approaching the maximum f -value of 1.00) for the SPM-test score on intelligence.

The cross-sequential results, including the absolute mean differences between the five cohorts (cross-sectional data) as well as these between the four times of measurement (longitudinal data), are shown graphically in Figure 2 for ideational fluency, Figure 3 for ideational flexibility, and Figure 4 for intelligence. At first glance the good overlap of cross-sectional and longitudinal developmental gradients is visible for both indicators of divergent and the indicator of convergent thinking skills. This illustrates and confirms the agreement of the ANOVA main effects of Cohort and Time of Measurement. The cross-sequential results on intelligence (reasoning; see Figure 4) show a very continuous increase from the next to last kindergarten year to the 6th year of elementary school. There are no obvious developmental plateaus, accelerations or delays, increases or decreases.

Table 2

Intercorrelations of Ideational Fluency (FLU), Ideational Flexibility (FLE), and Intelligence (IN) at (in Bold) and Between the Four Times of Measurement in the Luxembourg Sample (Above the Main Diagonal) and in the German Sample (Below the Main Diagonal)

Variable	T1			T2			T3			T4		
	FLU	FLE	IN	FLU	FLE	IN	FLU	FLE	IN	FLU	FLE	IN
1st measurement (T1)												
FLU	1.00	.64**	.19**	.49**	.41**	.08	.32**	.23**	.10	.30**	.16*	.11
FLE	.60**	1.00	.26**	.36**	.49**	.16*	.31**	.31**	.17**	.28**	.28**	.14*
IN	.14*	.24**	1.00	.09	.18**	.64**	.10	.15*	.54**	.08	.12	.48**
2nd measurement (T2)												
FLU	.44**	.30**	.11	1.00	.65**	.27**	.48**	.52**	.22**	.29**	.24**	.07
FLE	.47**	.44**	.16**	.61**	1.00	.28**	.35**	.45**	.11	.28**	.27**	.12
IN	.71**	.20**	.66**	.19**	.23**	1.00	.05	.10	.53**	.11	.10	.59**
3rd measurement (T3)												
FLU	.39**	.27**	.07	.45**	.36**	.07	1.00	.59**	.21**	.27**	.31**	.12
FLE	.36**	.32**	.17**	.45**	.43**	.15**	.54**	1.00	.26**	.30**	.27**	.14*
IN	.65**	.21**	.55**	.60**	.09	.64**	.19**	.24**	1.00	.27**	.14*	.61**
4th measurement (T4)												
FLU	.31**	.26**	.10	.25**	.29**	.10	.25**	.28**	.26**	1.00	.55**	.20**
FLE	.30**	.15**	.06	.25**	.28**	.14*	.22**	.23**	.23**	.48**	1.00	.27**
IN	.62**	.12*	.51**	.67**	.03	.59**	.69**	.21**	.65**	.18**	.31**	1.00

* $p < .05$. ** $p < .01$.

Table 3

Results of Analyses of Variance for Ideational Fluency (FLU), Ideational Flexibility (FLE), and Intelligence (IN) With the Between Factor Educational Cohort (5) and the Repeated Measurement Factor Time of Measurement (4) in the Luxembourg Sample ($N = 228$) and the German Sample ($N = 294$)

MANOVA effect	Ideational fluency	Ideational flexibility	Intelligence
Luxembourg students			
Cohort			
F(4/224)	25.35**	28.60**	36.57**
Effect size (f)	.53	.61	.77
Time			
F(3/672)	33.82**	37.49**	59.44**
Effect size (f)	.44	.57	.89
Cohort x Time			
F(12/672)	6.80	4.07	2.89
Effect size (f)	.06	.09	.01
German students			
Cohort			
F(4/290)	27.18**	27.90**	38.11**
Effect size (f)	.59	.60	.79
Time			
F(3/870)	49.05**	38.89**	61.31**
Effect size (f)	.67	.65	.92
Cohort x Time			
F(12/870)	7.42	5.79	3.01
Effect size (f)	.05	.05	.02

** $p < .01$.

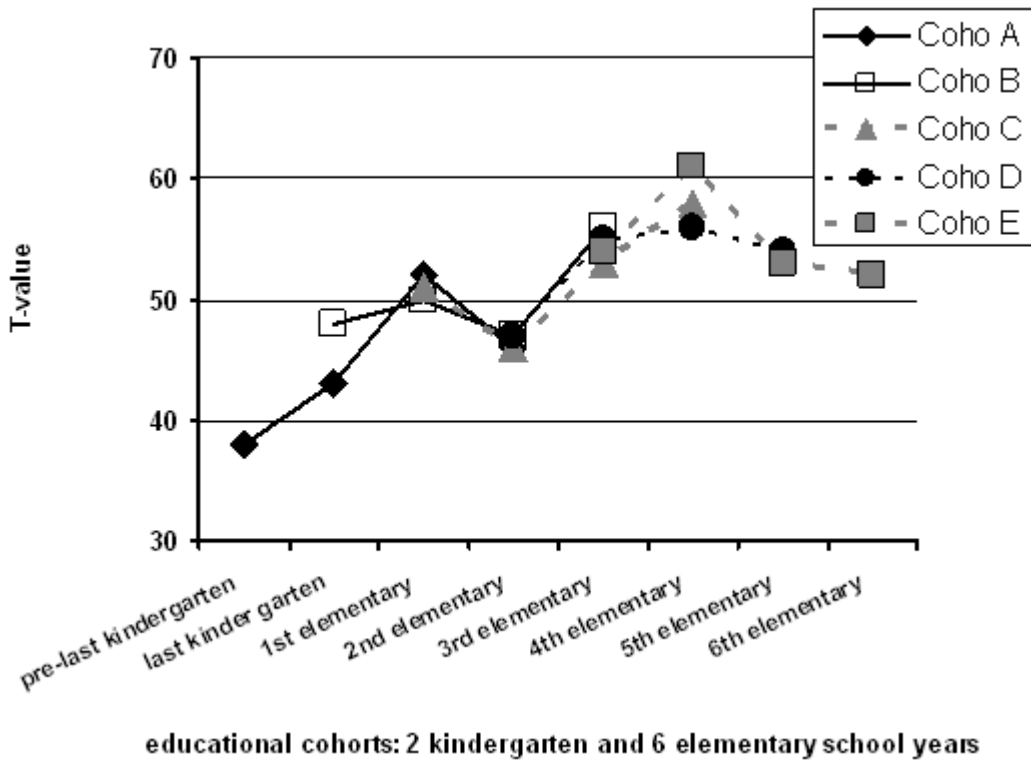


Figure 2. Cross-sequential results on ideational fluency in the Luxembourg sample.

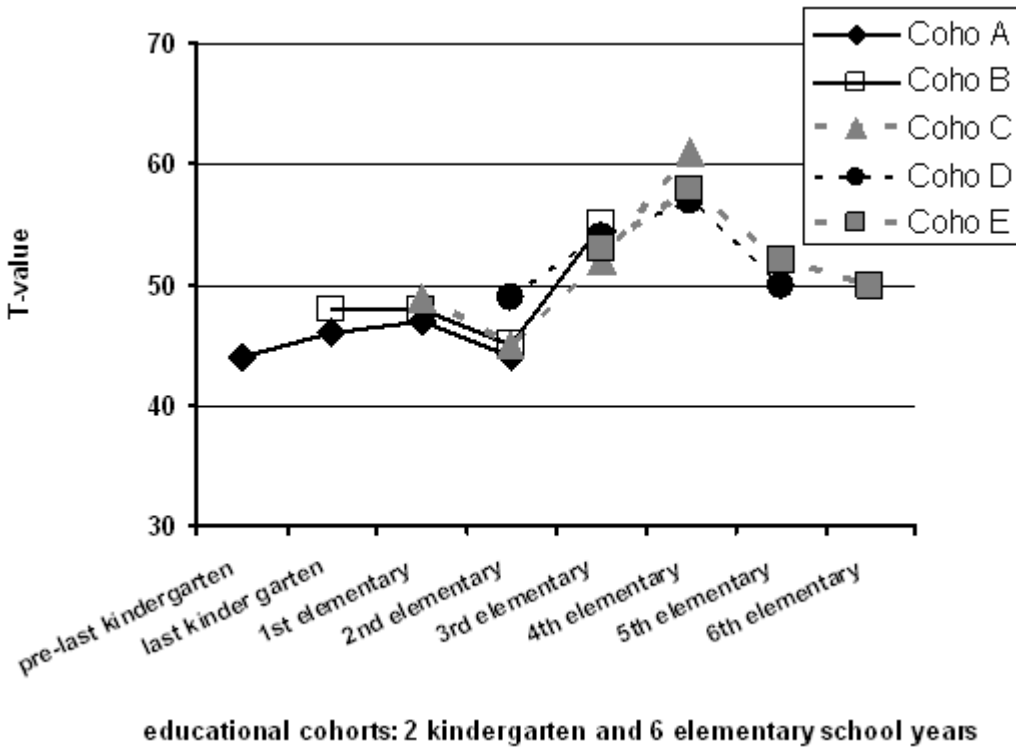


Figure 3. Cross-sequential results on ideational flexibility in the Luxembourg sample.

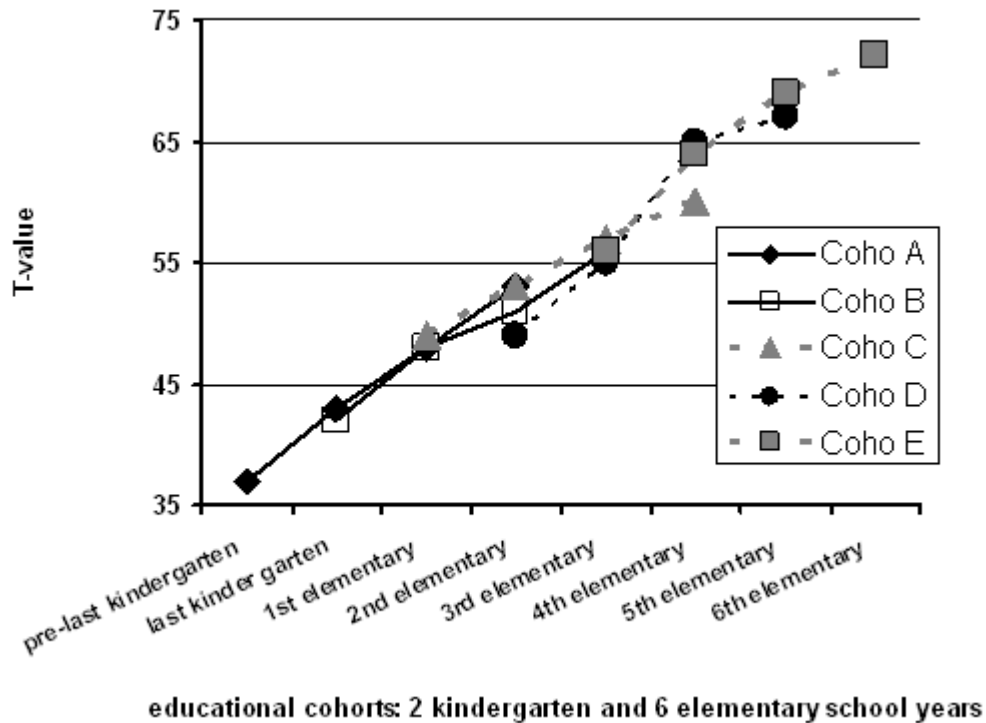


Figure 4. Cross-sequential results on general intelligence in the Luxembourg sample.

In contrast, both indicators of creativity show marked, discontinuous developmental gradients: After a steady increase, ideational fluency drops from the 1st to the 2nd grade level of elementary school and, again, from the 4th to the 5th and 6th grade level (see Figure 2). Evaluated by post hoc single mean comparison tests (Scheffé test), both declines of fluency are statistically significant ($p < .05$). Very similar declines are present in the transitions of the 1st to the 2nd and of the 4th to the 5th/6th grade level in the Luxembourg elementary school children for ideational flexibility (see Figure 3). Post hoc single mean comparisons confirm the statistical significance for both transitions points ($p < .05$).

Short Discussion of Study I: Luxembourg

A double “creativity slump” is apparent in the Luxembourg sample of children who are schooled in an educational system with obligatory two years of kindergarten and 6-year comprehensive elementary school education: Longitudinal and cross-sectional data are in full agreement for a first decline in ideational fluency and flexibility in the transition from the 1st to the 2nd and a second decline in the transition from the 4th to the 5th/6th elementary school grade level, which is connected to the anticipation of educational placement for secondary education after the 6th elementary school grade level. The soon coming selective educational placement is anticipated by the children, parents, and teachers in the final two years of elementary school education. The anticipation of educational placement is waiting to be dealt with, it is on the agenda of parents, teachers, and children. This results in more or less stress, because (convergent) academic achievement and educational degrees are strongly in the focus. In contrast, the development of intelligence is increasing very continuously. Positional stability of intelligence score is high, positional stability of scores on divergent thinking is much lower. This points towards large interindividual differences in the intraindividual development of creativity in childhood. Cross-sectional and longitudinal data

are—with reference to reliability and scoring inter-coder reliability as well as to the lack of hints on retest effects and dropout biases—sound. However, the results refer to only one specific educational administration system, and therefore, the question of whether similar, other, or no “creativity slumps” can be seen in another educational system with different transition points is still open.

Study II: The Germany Study

Method

Sample — Participants of the study were 312 kindergarten and elementary school students (some of them becoming secondary school students at the later times of measurement) living in Southwest Germany in the city of Trier (approximately 100,000 inhabitants) as well as in surrounding towns and villages. This German region is located in the heart of Western Europe and is situated directly next to the open European border to Luxembourg. Children were selected randomly (controlling for sex) from five different educational cohorts: Children were enrolled in the next to last kindergarten year ($n = 62$), last kindergarten year ($n = 62$) as well as in the 1st ($n = 63$), 2nd ($n = 62$), and 3rd elementary school year ($n = 63$). There were three refusals for participation, two from parents and one from a kindergarten child. Age span at initial test was 4 to 9 years ($M = 6.8$, $SD = 1.61$ years). The sample comprised 156 females and 156 males.

Just as in the Luxembourg study, the children were retested one year as well as two and three years after the initial test (T1). All tests were administered in the winter season after being in the (new) educational cohort, i.e., kindergarten year, elementary school or secondary school grade level, for approximately 10 weeks. Thus, longitudinal data refer to four times of measurement within three years (T1, T2, T3, and T4). This builds up a cross-sequential design with 5 educational cohorts and 4 times of measurement just as in Study I. As a consequence, at T2 children were enrolled to the last kindergarten year up to the 4th elementary school year, at T3 to the 1st elementary up to the 5th school year (i.e., the 1st grade level of secondary education), and at T4 data cover up the 2nd elementary school year up to the 6th grade level (i.e., the 2nd grade level of secondary education). Age span at T4 was T1-age plus 3 years, i.e., 7 to 12 years ($M = 9.7$, $SD = 1.62$). Because of continuous contact with the children and their teachers, dropout rates were small: Retests could be administered at T2 with 310 (99%), at T3 with 307 (98%), and at T4 with 295 children (95%) of the initial sample tested at T1. Dropout analyses did not indicate any biases with reference to the educational cohorts and to the dependent variables tested at the time(s) of measurement before.

In addition, for the control of possible retest effects in test scores, at T2 $N = 90$ children (enrolled in the last kindergarten year and 1st to 4th elementary school year) were tested for the first time.

Measures — The same culture-free, predominantly nonverbal tests on divergent and convergent thinking as well as scoring procedures as in the Luxembourg study (see Section: Luxembourg Study) were administered in individual testing.

Inter-coder reliability of the *FLU-* and *FLE-*scorings was tested by comparisons of independent scorings for the records of the total sample tested at T1 ($N = 312$) and of randomly selected subsamples of $n = 40$ tested each at T2 and T3, respectively. Interscorer agreement was confirmed by the correlations between two independent scorers for FLU ($r > .96$; $p < .01$) and FLE ($r > .89$; $p < .01$) and by the nonsignificant mean differences between

the two independent scorings of FLU ($t(310) = 1.53$ and $t(38) < 0.63$; $p > .20$) and of FLE ($t(310) = 0.97$ and $t(38) < 0.77$; $p > .20$).

Reliability of the FLU- and FLE-scores is evaluated by Cronbach's alpha for each of the four times of measurement (see Table 4). The reliability coefficients of FLU and FLE are satisfactory for the intended statistical analyses at the group level (FLU: $\alpha > .82$ and FLE $\alpha > .73$). Inter-coder reliability of the SPM, analyzed by two independent scorers as above, in the present study is $r = 1.00$ without any mean differences ($t(xx) = 0.00$). Reliability coefficients for all of the four times of measurement are very good ranging from .89 to .93 (see Table 1).

Table 4

Coefficients of Internal Consistency (Cronbach's α) at the Four Times of Measurement and Coefficients of Test-Retest Stability Between the Four Times of Measurement in the German Sample

Time of measurement	N	Scale		
		Fluency	Flexibility	Intelligence
Internal consistency (α)				
1 st measurement (T1)	312	.83	.75	.90
2 nd measurement (T2)	310	.87	.74	.93
3 rd measurement (T3)	307	.90	.78	.91
4 th measurement (T4)	295	.91	.77	.90
Test-retest-stability (r)				
T1 - T2	310	.44**	.47**	.71**
T1 - T3	305	.39**	.36**	.65**
T1 - T4	294	.31**	.30**	.62**
T2 - T3	307	.45**	.40**	.60**
T2 - T4	294	.25**	.25**	.67**
T3 - T4	294	.25**	.22**	.69**

** $p < .01$.

Data on the total *test duration* and on personal characteristics of the child (e.g., age, sex, family size) were also documented in the test record. Test time varied between $M = 62$ ($SD = 7.1$) minutes at T1, $M = 60$ ($SD = 6.9$; T2), $M = 61$ ($SD = 6.3$; T3) and $M = 62$ ($SD = 6.5$; T4) without significant differences between the repeated measurements [$F(3/870) = 1.84$]. There are nonsignificant ANOVA effects of the educational cohort [$F(4/290) = 0.89$] and its interaction with time of measurement [$F(12/870) = 2.55$]. There are significant, but rather low correlations of test time with the test scores on ideational fluency (FLU: $.19 < r < .35$; $p < .01$) and ideational flexibility (FLU: $.17 < r < .20$; $p < .01$), but only one significant correlation out of four correlations with the SPM-scores on intelligence (T4: $r = .20$, $p < .01$; T1, T2, T3: $r < .08$). Thus, common variance of the scores on divergent thinking with test duration is at maximum 12%, and the common variance of the scores on convergent thinking with test duration is at maximum 4%.

Results of Study II: Germany

Test-Retest Effects — Checks for retest effects on test scores use the data from the $N = 90$ children (enrolled in the last kindergarten year and 1st to 4th elementary school year) tested for their first time at T2. Comparisons of this new sample with the longitudinal sample retested at T2 did not result in statistical significant mean differences for neither the test scores on ideational fluency [$t(400) = 0.83$; $p > .10$] and flexibility [$t(400) = 1.04$; $p > .10$] nor for intelligence [$t(400) = 0.75$; $p > .10$].

Positional Stability Versus Plasticity Over Time — Relative (i.e., positional) changes examined longitudinally in the 4-year time span are low to medium in ideational fluency and flexibility, but they are markedly higher in intelligence (see Table 4). Correlation coefficients of FLU- and FLE-scores, respectively, decline to less than 6% common variance between more distant times of measurement, while common variance of SPM-scores between all different times of measurement exceed 36% and range up to 50%. Thus, high positional stability is confirmed for culture-free tested convergent thinking skills from kindergarten up to the 6th grade level of school education in Germany. On the other hand, positional stability of culture-free tested divergent thinking skills in childhood development tested one to three years apart is low.

Intercorrelations of Divergent and Convergent Thinking Scores at and Between Times of Measurement — Intercorrelations of ideational fluency and flexibility as well as intelligence at the four times of measurement and between them are presented below the main diagonal in Table 2 for the German sample. Cross-sectional correlations (computed with the data from one time of measurement, printed in bold in Table 2) are higher than the longitudinal correlations (computed with the data of two different times of measurement). This difference is, on average, statistically significant at $p < .05$. The two nonverbal scales on divergent thinking (FLU and FLE) correlate at a statistically significant level, with a common variance up to 37%. The nonverbal test score on convergent thinking (SPM), however, is correlated much lower with FLU and FLE, and the correlations lose in part their significance in time-lagged analyses (see Table 2). The trend seems to indicate somewhat slightly higher correlations of general intelligence (reasoning) with ideational flexibility (FLE) by comparison to ideational fluency, but the common variance of culture-free, predominant nonverbally tested creativity and intelligence in children does not exceed 10% in the German sample.

Absolute Changes (Mean-Level): Cohort and Time Effects — Just as in the Luxembourg Study, cross-sequential data analyses were computed by analyses of variance (ANOVA) with the between-factor *Five Educational Cohorts* and the repeated measurement factor of *Four Times of Measurement*, each being one year apart from the other. Results of the computed ANOVAs are presented for the German sample in the lower half of Table 3. There are no statistically significant interaction terms. The main effects of the between (Cohort) and repeated measurement (Time) factors are statistically significant. Following the terminology of Cohen (1977), the effect sizes (f) of all ANOVA main effects are large.

The cross-sequential results are presented graphically in Figures 5 to 7. The overlap of cross-sectional and longitudinal developmental gradients is high for both indicators of divergent and the indicator of convergent thinking skills. This illustrates and confirms the agreements of the ANOVA main effects of Cohort and Time of Measurement. The cross-sequential results on intelligence (reasoning; see Figure 7) show a continuous, very steady increase from the next to last kindergarten year to the 6th grade school level. There are no developmental plateaus, accelerations or delays, increases or decreases.

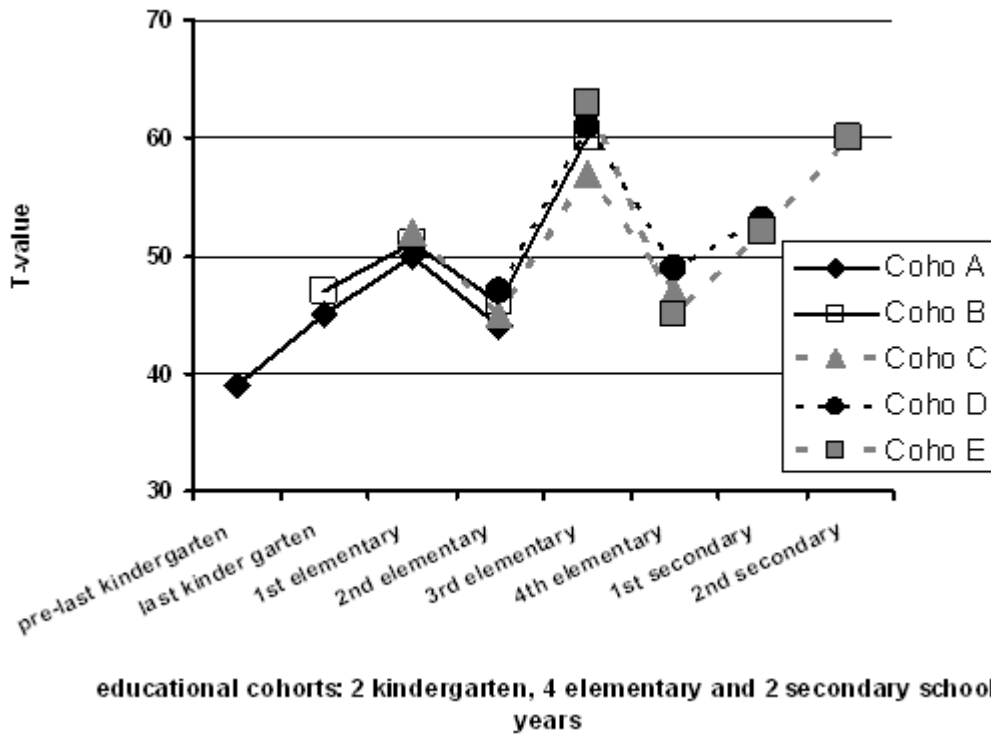


Figure 5. Cross-sequential results on ideational fluency in the German sample.

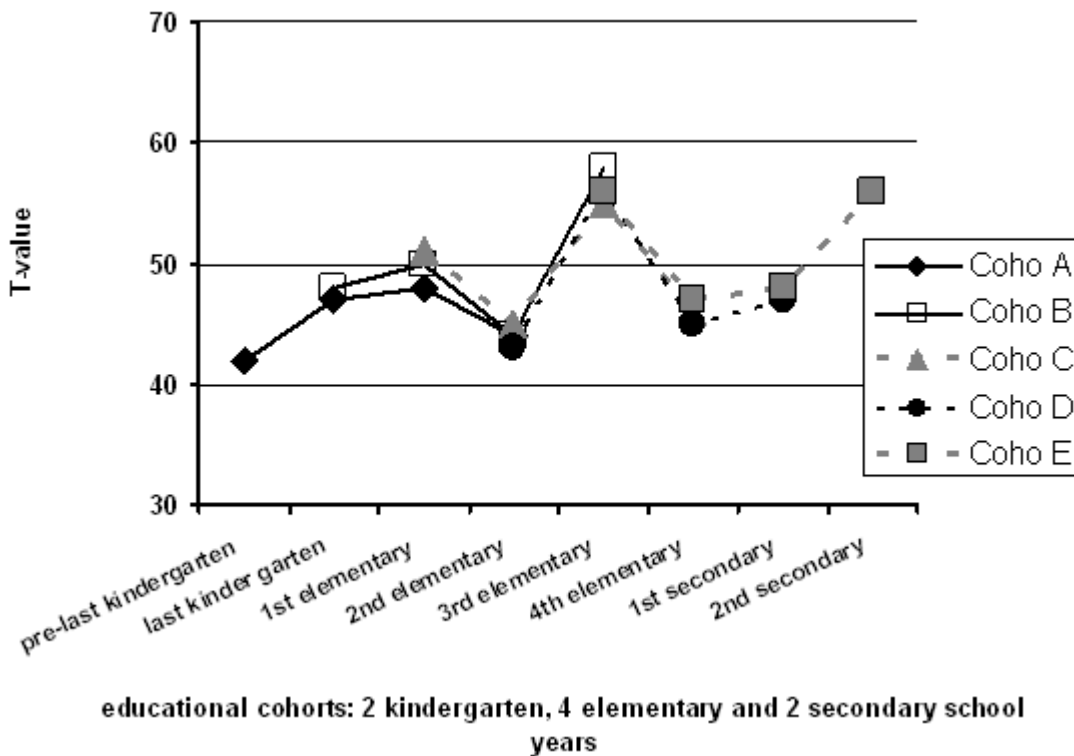


Figure 6. Cross-sequential results on ideational flexibility in the German sample.

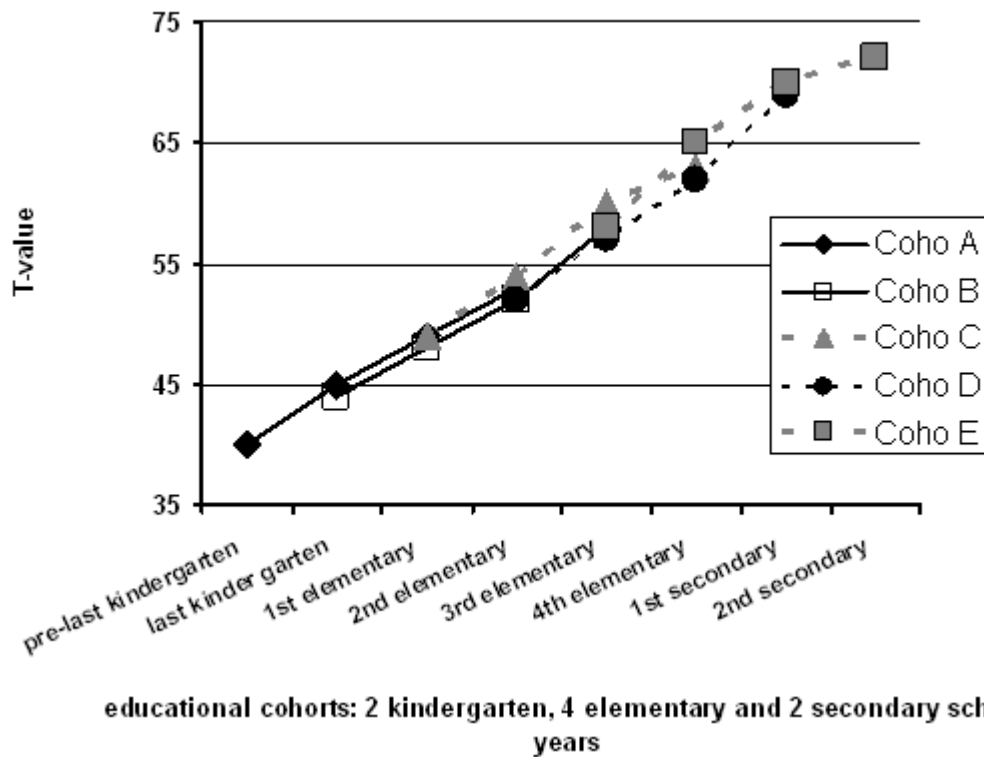


Figure 7. Cross-sequential results on general intelligence in the German sample.

In contrast, both indicators of creativity show marked discontinuous developmental gradients: After an increase, ideational fluency drops from the 1st to the 2nd grade level of elementary school and, after a strong increase in the 3rd grade level, there is a second drop from the 3rd to the 4th grade level and a steady increase thereafter (see Figure 2). Evaluated by post hoc single mean comparison tests (Scheffé test), both declines of fluency are statistically significant ($p < .05$). Very similar declines are present in the transitions of the 1st to the 2nd and of the 3rd to the 4th grade level in the German elementary school children for ideational flexibility (see Figure 3). Post hoc single mean comparison confirms the statistical significance for both transition points ($p < .05$).

Short Discussion of Study II: Germany

A double “creativity slump” is apparent in the German sample of children who are schooled in an educational system with optional kindergarten and a 4-year comprehensive elementary school education, which is followed by educational placement into secondary education: Longitudinal and cross-sectional data are in agreement for a first decline in ideational fluency and flexibility in the transition from the 1st to the 2nd school year and a second decline during anticipation of educational placement after the 4th school grade level in the transition from the 3rd to the 4th elementary school grade level. In contrast, development of intelligence is very continuous. Positional stability of intelligence scores is high, positional stability of scores on divergent thinking is lower. Cross-sectional and longitudinal data are—with reference to reliability and inter-coder reliability as well as to the lack of hints on retest effects and dropout biases—sound.

General Discussion

The cross-sequential developmental results presented confirm significant “creativity slumps” in the domain of ideational fluency and flexibility in strong connection to educational transitions and—in older children—to the anticipation of transitions by children, their parents, and teachers, findings which are significant for educational placement (here, e.g., educational placement into secondary educational tracks with consequences for the educational and future occupational options of the children). These results are convincing, and more importantly, they are a powerful empirical confirmation of the hypothesis of [Torrance \(1963\)](#) (1) because high standards of developmental psychology research are assured (i.e., cross-sequential design and control of retest effects in repeated measurements), (2) because culture-free, predominantly nonverbal psychometric test indicators of divergent and convergent thinking are administered in individual testing (i.e., thus to a large extent avoiding confounding biases with reference to language development and social status of the children), and (3) because a quasi-experimental approach was implemented investigating the given reality of two different national educational administration systems in neighboring yet distinct European regions. The fact that the reliability of the tests on divergent thinking are slightly lower than that of the intelligence measurement applied can be neglected, because the differences between the reliability coefficients are statistically not significant ($p > .20$; [Feldt, Woodruff, & Salih, 1987](#)).

However, it must be noted that the quasi-experimental design was feasible only indirectly. This was necessary because of differences in test administration in a multilingual (Luxembourg) *versus* monolingual (Germany) society, nation, and educational system. Thus, the quasi-experimental approach resulted in convincing descriptive empirical confirmations of declines in divergent thinking after and before anticipated educational transitions which are strongly connected to specifics of the national educational administration systems under study. However, exploratory power of the results remains low, because multilingual test administration in the Luxembourg sample and monolingual testing in the German one required separate data analyses including within-sample test score standardizations (within sample standardization resulting in T-values; [McCall, 1939](#)). This excluded the possibility of direct statistical tests of the described developmental differences between Luxembourg and German kindergarten and school children. Further on, the identity of the two countries is kept and in the cross-national comparisons each national context presents some specificity in terms of language and educational system.

At the very least, the results point to the fact that [McCrae et al.'s \(1987\)](#) findings that cross-sectional and longitudinal (repeated measures) analyses suggesting curvilinear trends in divergent thinking in adult men do not hold up in cross-sequential analyses (which point to a general, rather linear decline in divergent thinking for all cohorts aged initially 33 to 74 years tested six years later) cannot be generalized to childhood development in educational contexts. The presented results of two independent studies agree completely in the confirmation of creativity slumps in childhood development which have only been empirically described before in cross-sectional (e.g., [Jaarsveld et al., 2010](#); [Klausmeier & Wiersma, 1964](#); [Pickard, 1979](#); [Smith & Carlsson, 1983](#); [Torrance, 1968](#)) or longitudinal studies (e.g., [Claxton et al., 2005](#); [Gaspar, 2001](#); [Howieson, 1981](#); [Torrance, 1968, 1981a](#)). In the focus of the present studies are critical educational transitions and their anticipation by children, their parent, and teachers in face of important educational placements.

Empirical evidence for a first decline of divergent skills is given for school enrollment after kindergarten education, which can be attributed to the rising demands on children's social adjustment, conformity, and acceptance of authorities as well as their requests for convergent thinking styles, which is frequently thought to be necessary

for the learning of cultural skills like reading, writing, mathematics, etc. However, these (traditional) demands are dependent of educational theory, teaching methods, and didactics applied in elementary education. For example, cross-sectional results of [Besançon and Lubart \(2008\)](#) and longitudinal results of [Heise et al. \(2010\)](#) confirm empirically more positive developments of creativity in Montessori Schools in comparison to traditional elementary schools. Montessori school education focuses on discovery teaching methods, individualized learning plans, and self-regulated learning, which is in contrast to traditional elementary school education characterized by more directed teaching methods in small group learning and teacher-centered teaching with general learning objectives for all students in the class, etc. being the widespread educational reality in Luxembourg and Germany. However, both countries were struck by the so-called “PISA and TIMSS shock” in the 1990s, because both had to face the fact that their elementary and secondary school education rankings in worldwide comparisons of school achievement were quite low (see, e.g., [Masters, 2007](#); Organization for Economic Cooperation and Development - OECD, 2009; [Rindermann, 2007](#)). This shock motivated both public and political discussions of reforms in the national teacher education, teaching methods, and educational administration in favor of educational systems and individualized teaching methods administered in countries like Finland, Canada, and New Zealand. This discussion is ongoing to date without any convincing, wide-spread educational reforms taking place, because these are hindered by very different political positions and—in Germany—by the federal structure. In addition, these public and political controversies refer to the early educational placement strategies after the 4th elementary grade level in Germany and after the 6th grade level in Luxembourg as well. There have been some attempts and educational experiments to extend comprehensive education (at least in Germany), but there is no political and public consensus about this as well. This refers immediately to the second decline of creativity described in the two studies presented above: However, this second creativity slump is located before educational placement in secondary education, which is anticipated by the children and their parents and teachers in Germany at the 3rd grade level and in Luxembourg at the 5th and 6th grade level.

Both creativity slumps refer to the overstressing convergent, general teaching methods as thought to be only necessary in teaching culture skills like reading, writing, mathematics, etc., which can be taught and learned in more divergent educational settings added to the teaching methods as well (see, e.g., [Kaufman & Beghetto, 2009](#); [Krampen & Freilinger, 1996](#); [Renzulli & Reis, 2001](#)). However, “creativity in the schools is sometimes seen as a footnote, afterthought, or as an extra-curricular activity (...)” ([Kaufman & Beghetto, 2009](#), p. 175). Beginning in the 1960s, demands for more creativity promotion in schools by divergent teaching methods, creativity techniques, creativity-promoting educational settings, and school/classroom climate (see, e.g., [Getzels & Elkins, 1964](#); [Sternberg & Lubart, 1991](#); [Torrance, 1963, 1968, 1981a](#)) must be repeated incessantly until they finally are taken into consideration in educational politics, educational administration, society, and schools.

At least, the results of some multilevel model analyses on the predictive power of creativity indicators for general academic achievement provide some hope: [Freund and Holling \(2008\)](#) as well as [Gralewski and Karwowski \(2012\)](#) showed empirically that in some schools or classrooms, respectively, the relations of creativity to scholastic achievement were positive and statistically significant, while in others (the majority) they were negative or insignificant. Gralewski and Karwowski’s results show that the role of creativity is greater in larger schools and in these located in big cities, Freund and Holling attribute their result that the predictive power of creativity changes across classrooms to the hypotheses that some teachers value creativity in their students more than others. Interindividual differences in teachers’ evaluation of students’ creativity and their teaching methods may be one reason for these findings, other reasons may refer to interindividual differences in parents’ educational attitudes and/or students’ characteristics, more liberal educational and more innovative attitudes in larger cities, etc. as

well. However, it must be considered that both studies are cross-sectional, neither longitudinal nor cross-sequential and therefore the interpretation of their results is limited, because there is no consecutive testing of creativity and academic performances in longer time intervals.

Divergent thinking and action of children must be seen consensually as highly significant for their own personal development and for societal development as well. The de-evaluation of creativity and its forerunners in divergent skills (see Figure 1) as not relevant, as not appropriate to educational practice, and as viewed to be potentially disruptive and negative must be overcome widely in educational politics, educational administration, society, and schools, e.g., in the majority of teachers, parents, and students. Perhaps the spreading of positive psychology in schools (Furlong, Gilman, & Huebner, 2009) can be a revival of the older demands. The cross-sequential developmental results presented here complete these argumentations empirically with reference to high developmental research standards.

Last, but not least, two other patterns of results should be mentioned. Firstly, our 4-year longitudinal data on the development of divergent thinking skills confirm prior results on the low positional stability of ideational fluency and flexibility at kindergarten and elementary school age (e.g., Gaspar, 2001; Magnusson & Backteman, 1978; Sparfeldt et al., 2009). This is in sharp contrast to the very high positional stability coefficients observed for intelligence and indicates the developmental plasticity of creativity development in children that gives a lot of space and possibilities to promote positive individual development. Secondly, our results from two independent studies confirm prior results on low correlations of divergent and convergent thinking skills, at least, when both are measured in culture-free, predominantly nonverbal ways (Jaarsveld et al., 2010; Wallach & Kogan, 1965; Wechsler et al., 2010). Thus, the general hypothesis about the existence of significant correlations between intelligence and creativity (or only below a certain intelligence threshold) may be true for their verbal subfacets and measurements (Getzels & Jackson, 1962; Hasan & Butcher, 1966; Magnusson & Backteman, 1978; Schubert, 1973). It should be considered that both verbal subfacets are biased by language development and social status and, therefore, may be interrelated *a priori* in the domain of crystallized intelligence. Thus, culture-free, predominantly nonverbal testing of divergent and convergent thinking skills is not only an optimization of test fairness (especially for the disadvantaged), but it also points to the possibilities of creativity promotion in all children independently of their intelligence level.

Funding

This research project was supported by the Ministère de l'Éducation et de la Jeunesse/Luxembourg.

Acknowledgments

Thanks to Dr. Joseph Freiling, Dr. Louis Wilmes, Joseph Medernach and the teacher students at the former "Institut Supérieur d'Études et de Recherches Pédagogiques" in Walferdange (Luxembourg) as well as the psychology students at the University of Trier (Germany) for their valuable support in the data collection.

References

- Acharyulu, S. T., & Yasodhara, P. (1984). Assessment of creative thinking abilities of preschool children through spontaneous drawings. *Psychological Studies*, 29, 192-196.
- Ben Baruch, E., Bruno, J. E., & Horn, L. L. (1985). Development of time horizon perceptions in students. *Journal of the Association for the Study of Perception*, 20, 5-16.

- Besançon, M., & Lubart, T. (2008). Differences in the development of creative competencies in children schooled in diverse learning environments. *Learning and Individual Differences, 18*, 381-389. doi:10.1016/j.lindif.2007.11.009
- Claxton, A. F., Pannells, T. C., & Rhoads, P. A. (2005). Developmental trends in the creativity of school-age children. *Creativity Research Journal, 17*, 327-335. doi:10.1207/s15326934crj1704_4
- Cohen, R. (1977). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.
- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement, 11*, 93-103. doi:10.1177/014662168701100107
- Freund, P. A., & Holling, H. (2008). Creativity in the classroom: A multilevel analysis investigating the impact of creativity and reasoning ability on GPA. *Creativity Research Journal, 20*, 309-318. doi:10.1080/10400410802278776
- Furlong, M. J., Gilman, R., & Huebner, E. S. (Eds.). (2009). *Handbook of positive psychology in schools*. New York, NY: Routledge.
- Gaspar, M. (2001). Interindividual stability of personality traits of children aged 9-14. *Review of Psychology, 8*, 49-59.
- Getzels, J. W., & Elkins, K. (1964). Perceptual and cognitive development. *Review of Educational Research, 34*, 559-573.
- Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence*. New York, NY: Wiley.
- Gralewski, J., & Karwowski, M. (2012). (in press). Creativity and school grades: A case from Poland. *Thinking Skills and Creativity*. doi:10.1016/j.tsc.2012.03.002
- Guilford, J. P. (1964). The nature of creativity. In C. W. Taylor (Ed.), *Frontiers in psychology* (pp. 125-131). Chicago, IL: Scott, Foresman & Co.
- Guilford, J. P. (1950). Creativity. *The American Psychologist, 5*, 444-454. doi:10.1037/h0063487
- Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill.
- Hasan, P., & Butcher, H. J. (1966). Creativity and intelligence: A partial replication with scottish children of getzels' and jackson's study. *British Journal of Psychology, 57*, 129-135. doi:10.1111/j.2044-8295.1966.tb01013.x
- Heise, E., Böhme, E., & Körner, S. B. (2010). Montessori-orientierter und traditoneller Grundschulunterricht [Montessori education versus traditional education in elementary school]. *Psychologie in Erziehung und Unterricht, 57*, 273-289.
- Howieson, N. (1981). A longitudinal study of creativity — 1965-1975. *The Journal of Creative Behavior, 15*, 117-134. doi:10.1002/j.2162-6057.1981.tb00284.x
- Jaarsveld, S., Lachmann, T., Hamel, R., & van Leeuwen, C. (2010). Solving and creating Raven Progressive Matrices: Reasoning in well- and ill-defined problem spaces. *Creativity Research Journal, 22*, 304-319. doi:10.1080/10400419.2010.503541
- Kaufman, J. C., & Baer, J. (Eds.). (2006). *Creativity and reason in cognitive development*. New York, NY: Cambridge University Press.

- Kaufman, J. C., & Beghetto, R. A. (2009). Creativity in schools: A rapidly developing area of positive psychology. In M. J. Furlong, R. Gilman, & E. S. Huebner (Eds.), *Handbook of positive psychology in schools* (pp. 175-188). New York, NY: Routledge.
- Klausmeier, H. J., & Wiersma, W. (1964). Relationship of sex, grade level, and locale to performance of high IQ students on divergent thinking tests. *Journal of Educational Psychology, 55*, 114-119. doi:10.1037/h0041717
- Kogan, N., & Pankove, E. (1972). Creative ability over a five-year span. *Child Development, 43*, 427-442. doi:10.2307/1127546
- Krampen, G. (1996). *Kreativitätstest für Vorschul- und Schulkinder* [Creativity test for kindergarten and elementary school students]. Göttingen (Germany): Hogrefe.
- Krampen, G., & Freilinger, J. (1996). *Kreativitätsförderung in der Primarschule* [Creativity promotion in elementary schools]. Walferdange (Luxembourg): ISERP Press.
- Krampen, G., Seiger, C., & Steinebach, C. (2012). Kreativität [Creativity]. In R. Zihlmann, D. Jungo, & C. Steinebach (Eds.), *Praxis der Positiven Psychologie* [Applications of positive psychology] (pp. 76-84). Weinheim (Germany): Beltz.
- Magnusson, D., & Backteman, G. (1978). Longitudinal stability of person characteristics: Intelligence and creativity. *Applied Psychological Measurement, 2*, 481-490. doi:10.1177/014662167800200402
- Mainberger, U. (1977). *Test zum divergenten Denken (Kreativität)* [Test on Divergent Thinking (Creativity)]. Weinheim (Germany): Beltz.
- Masters, G. N. (2007). Introduction to the special issue on PISA. *Journal of Applied Measurement, 8*, 235-236.
- McCall, W. A. (1939). *Measurement*. New York, NY: Wiley.
- McCrae, R. R., Arenberg, D., & Costa, P. T. (1987). Declines in divergent thinking with age: Cross-sectional, longitudinal, and cross-sequential analyses. *Psychology and Aging, 2*, 130-137. doi:10.1037/0882-7974.2.2.130
- Organization for Economic Co-operation and Development (OECD). (2009). *Education at a glance*. Paris: OECD.
- Paris, C., Edwards, N., Sheffield, E., Mutinsky, M., Olexa, T., Reilly, S., & Baer, J. (2006). How early school experiences impact creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity and reason in cognitive development* (pp. 333-350). New York, NY: Cambridge University Press.
- Peterson, C., & Seligman, M. E. P. (2004). *Character strength and virtues: A handbook and classification*. Washington, DC: American Psychological Association.
- Pickard, E. (1979). *The development of creative ability*. Windsor (GB): NFER Publishing.
- Rattat, A. C., & Droit-Volet, S. (2005). Long-term memory for duration: Functioning and development. *Psychologie Française, 50*, 99-116. doi:10.1016/j.psfr.2004.10.006
- Raven, J. C., Court, J., & Raven, J., Jr. (1979). *Standard Progressive Matrices (SPM)*. Weinheim (Germany): Beltz.
- Renzulli, J. S., & Reis, S. M. (2001). Developing creativity in children through the enrichment triad model. In T. P. Gullotta & M. Bloom (Eds.), *Promoting creativity across the life span* (pp. 127-156). Washington, DC: Child Welfare League of America.

- Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21*, 667-706. doi:10.1002/per.634
- Schubert, D. S. P. (1973). Intelligence as necessary but not sufficient for creativity. *The Journal of Genetic Psychology, 122*, 45-47.
- Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *The American Psychologist, 55*, 5-14. doi:10.1037/0003-066X.55.1.5
- Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York, NY: Free Press.
- Simonton, D. K. (2000). Creativity. *The American Psychologist, 55*, 151-158. doi:10.1037/0003-066X.55.1.151
- Smith, G. J. W., & Carlsson, I. (1983). Creativity in early and middle school years. *International Journal of Behavioral Development, 6*, 167-195.
- Sparfeldt, J. R., Wirthwein, L., & Rost, D. H. (2009). Hochbegabt und einfallslos? [Gifted and not creative]. *Zeitschrift für Pädagogische Psychologie, 23*, 31-39. doi:10.1024/1010-0652.23.1.31
- Sternberg, R. J., & Lubart, T. I. (1991). An investment theory of creativity and its development. *Human Development, 34*, 1-31. doi:10.1159/000277029
- Torrance, E. P. (1963). *Education and the creative potential*. Minneapolis, MI: University of Minnesota Press.
- Torrance, E. P. (1968). A longitudinal examination of the fourth grade slump in creativity. *Gifted Child Quarterly, 12*, 195-199.
- Torrance, E. P. (1981a). Predicting the creativity of elementary school children (1958-80) and the teachers who "made a difference". [Reprint in D. J. Treffinger (Ed.), *Creativity and giftedness* (pp. 35-49). Thousand Oaks, CA: Corwin]. *Gifted Child Quarterly, 25*, 55-62. doi:10.1177/001698628102500203
- Torrance, E. P. (1981b). *Thinking Creatively in Action and Movement*. Bensenville, IL: Scholastic Testing Service.
- Treffinger, D. J. (1980). *Encouraging creative learning for the gifted and talented*. Ventura, CA: LTI Publications.
- Treffinger, D. T., Isaksen, S. G., & Firestien, R. L. (1983). Theoretical perspectives on creative learning and its facilitation: An overview. *The Journal of Creative Behavior, 17*, 9-17. doi:10.1002/j.2162-6057.1983.tb00970.x
- Wallach, M. A., & Kogan, N. (1965). A new look at the creativity-intelligence distinction. *Journal of Personality, 33*, 348-369. doi:10.1111/j.1467-6494.1965.tb01391.x
- Wechsler, S. M., Oliveira Nunes, M. F., Schelini, P. W., Ferreira, A. A., & Pascoal, P. A. (2010). Creativity and intelligence: Analyzing developmental similarities and discrepancies. *Estudos de Psicologia, 15*, 243-250. doi:10.1590/S1413-294X2010000300003
- Zachopoulou, E., Makri, A., & Pollatou, E. (2009). Evaluation of children's creativity: Psychometric properties of Torrance's 'Thinking Creatively in Action and Movement Test'. *Early Child Development and Care, 179*, 317-328. doi:10.1080/03004430601078669

About the Author

Günter Krampen is a Professor of Psychology at the University of Trier (Germany) and an Associate Professor of Psychology at the Université du Luxembourg (Luxembourg).