Two decades of research on children's creativity development during primary education in relation to task characteristics

Potters, O.T.A.; van Schijndel, T.J.P.; Jak, S.; Voogt, J.

DOI
10.1016/j.edurev.2023.100532

Publication date
2023

Document Version
Final published version

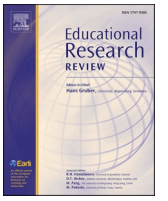
Published in
Educational Research Review

Educational Research Review 39 (2023) 100532

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev



Review

Two decades of research on children's creativity development during primary education in relation to task characteristics

Olga T.A. Potters [a,b,*], Tessa J.P. van Schijndel [a], Suzanne Jak [a], Joke Voogt [a]

[a] Research Institute Child Development and Education, University of Amsterdam, the Netherlands
[b] ArtEZ University of the Arts, Arnhem, the Netherlands

ARTICLE INFO

Keywords:
Creativity development
Primary education
Literature review
Meta-analysis
Task characteristics

ABSTRACT

This systematic literature review aimed to gain more insight into research on the development of children's creativity in primary education in relation to different task characteristics. The review introduces a taxonomy to map creativity tasks. The taxonomy is based on a broad conceptualization of creativity, and differentiates between open- and close-ended, and integrating and fragmenting creativity tasks. Besides using the taxonomy, we also analyzed tasks in terms of the used domains of the stimulus and response of the tasks. A total of 52 studies published in the last 20 years were systematically analyzed with a combination of descriptive analyses and meta-analyses. The body of research demonstrated a varied picture concerning the measurement of creativity and its subdimensions. Open-ended integrating and open-ended fragmenting tasks were frequently used. The results showed a general increase in creativity across primary education. Less pronounced developmental patterns were found for the originality subdimension (open-ended fragmenting taxonomy quadrant), and for general creativity and some subdimensions when tasks with a combination of figural and verbal responses were used. The taxonomy of creativity tasks, introduced in this review, can be useful for both clarifying discussions in the creativity research discourse and designing creativity activities for educational practice.

1. Introduction

There is consensus among scholars and policymakers about the need for education to prepare children for a changing and complex world. Flexibility and creative solutions are needed to face societal developments and problems (e.g., Fullan & Langworthy, 2014; Glăveanu & Kaufman, 2019). That is why creativity is seen as a key skill to be developed in education (e.g., Beghetto, 2016; Saywer, 2012), and understanding of creativity development in children is needed. This systematic literature review concerns creativity development in primary education.

Previous studies report different developmental patterns of creativity in the primary years: on the one hand positive developmental patterns are detected (e.g., Alfonso-Benlliure & Santos, 2016; Besançon & Lubart, 2008; Maker, Jo, & Muammar, 2008), but also irregular developmental patterns with sudden drops (slumps) associated with different childhood stages are found (e.g., Kim, 2011; Krampen, 2012; Torrance, 1968). Two slumps or plateaus in creativity development are frequently found: one between grade 3 and 4 (grade 4 slump; e.g. Darvishi & Pakdaman, 2012; Lubart & Lautrey, 1995; Lubart & Georgsdottir, 2004; Torrance, 1967, 1968), and one between grade 5 and 6 (grade 6 slump/plateau; e.g., Alacapinar, 2013; Gralewski, Gajda, Wisniewsska, Lebuda, & Jankowska,

* Corresponding author.
E-mail address: o.t.a.Potters@uva.nl (O.T.A. Potters).

https://doi.org/10.1016/j.edurev.2023.100532
Received 23 March 2022; Received in revised form 19 March 2023; Accepted 20 March 2023
Available online 27 March 2023

2016; Jastrzebska & Limont, 2017; Kim, 2011). In this review we focus on the task level to better understand creativity development during primary education. As in the field on creativity development a wide array of instruments containing different types of tasks is applied, a task-centered approach could provide an overview on how creativity is measured in developmental research and may possibly provide more clarity in the patterns of creativity development. Furthermore, the focus on creativity task characteristics can be considered promising for educational practice, as it can contribute to the development of educational interventions.

Most literature reviews and meta-analyses on creativity in the primary years do not focus on creativity development. For example, meta-analyses focus on creativity and academic achievement (Gajda, Karwowski, & Beghetto, 2017), on creativity and the learning environment (Davies et al., 2013), and on creativity training programs (Scott, Leritz, & Mumford, 2004). Two recent literature reviews (Kupers, Lehmann-Wermser, McPherson, & Van Geert, 2019; Preiss, Grau, Ortiz, & Bernardino, 2016), and one meta-analysis (Said-Metwaly, Fernández-Castilla, Kyndt, Van den Noortgate, & Barbot, 2020) do focus on children's creativity development in primary education, but two of the three studies do not relate patterns in creativity development to task characteristics, nor do these studies include kindergarten. Preiss et al. (2016) limit the scope of their review to studies that have been performed in South America. Kupers et al. (2019) aimed at integrating existing theories of creativity in a dynamic systems model of creativity, and subsequently used this model to systematically review the empirical literature on creativity. The authors give an overview of the quantity of research focused on creativity development, but do not discuss developmental trajectories of creativity. Said-Metwaly et al. (2020) operationalized creativity in their meta-analysis as divergent thinking and focused on developmental patterns in relation to some task characteristics of the applied instruments. The authors concluded that there was an upward trend in divergent thinking, and the existence of a fourth-grade slump was related to the use of a verbal task. In addition, an originality (sub-dimension of creativity) fourth-grade slump was observed when using the Wallach Kogan Creativity Test (WKCT; Wallach & Kogan, 1965), while a plateau in grade 4 was observed when using the Torrance Test of Creative Thinking (TTCT, Torrance, 1966) or other divergent thinking tests.

The current systematic literature review builds on the study of Said-Metwaly et al. (2020), and further investigates the development of creativity in the primary years in relation to task characteristics. To this end, we developed a new taxonomy of creativity tasks. While the meta-analysis of Said-Metwaly et al. purely focused on divergent thinking, the taxonomy in the current study captures a broad conceptualization of creativity. This choice is in line with advocacies for a broader operationalization of creativity (Barbot et al., 2019). Second, in the current study, developmental patterns are not only related to taxonomy-related task characteristics, but as creativity is considered (partly) domain-specific (e.g., Baer, 1998; Baer & Kaufman, 2005, 2017; Barbot, Besançon, & Lubart, 2016; Kaufman & Baer, 2005; Lubart & Georgsdottir, 2004), we also paid attention to the domain of the task. We add to previous work (e.g., Baer & Kaufman, 2005, 2017; Said-Metwaly et al., 2020; Zyga, Ivcevic, Hoffmann, & Palomera, 2021) by making a clear distinction between the domain of the stimulus, and the domain of the response of the task. Lastly, the current review includes kindergarten to provide a complete picture of research on children's creativity in the primary years.

2. Conceptualization of creativity

Plucker, Beghetto, and Dow (2004) conducted a content analysis of definitions of creativity in research. Based on this analysis, they proposed the following definition of creativity: "Creativity is the interaction among aptitude, process, and environment, by which an individual or group produces a perceptual product that is both novel and useful as defined within a social context" (p.90). That is, the two criteria to determine if something is creative are 'novel' and 'useful' (e.g., Plucker et al., 2004; Simonton & Damian, 2013; Sternberg & Lubart, 1999). On the one hand, a creative product is something novel, unexpected, and original, and goes beyond what is already known at a certain point. On the other hand, the product must also be appropriate and useful, it must provide a fitted solution to the task or problem given. Although these two characteristics of creativity are very different and might be considered opposite in nature, in creativity the combination of both is essential (e.g., Plucker et al., 2004; Simonton & Damian, 2013).

The two creativity criteria, novelty and usefulness, can be linked to two thinking processes: divergent and convergent thinking, first introduced by Guilford (1951, 1967). Divergent thinking is defined as thinking that explores various directions of thought (e.g., Guilford, 1967; Runco & Acar, 2012, 2019). Divergent thinking in creativity tests is mostly evaluated with the criteria originality, fluency, and flexibility (e.g., Guilford, 1950, 1967). Originality refers to the uniqueness or rarity of the relevant responses. Fluency refers to the amount of unique and relevant responses. Finally, flexibility refers to the different categories of the relevant responses. Convergent thinking, on the other hand, is defined as thinking that moves towards a single (or a few) best option(s). Convergent thinking in creativity refers to evaluating solutions to solve the problem on appropriateness (e.g., Cropley, 2006; Glăveanu, 2013, 2014). Furthermore, convergent thinking is useful when a conventional idea or a particular and correct solution is required. Divergent thinking and convergent thinking are in interplay in creativity (Guilford, 1967), and are not mutually exclusive (Runco & Acar, 2019). Both thinking processes are needed in making creative products (e.g., Cropley, 2006; Guilford, 1967; Runco, 2008; Simonton & Damian, 2013), and therefore included in every creative process. Even tasks which are usually described as divergent thinking tasks, need some convergent thinking as well and vice versa. For example, 'The unusual uses of a brick' task (Guilford, 1951), asks participants to come up with as many ideas as possible for which a brick can be used. The thinking processes related to this task are closely related to divergent thinking, because participants explore many directions of thought (e.g., Guilford, 1967; Runco & Acar, 2012, 2019). But this task asks for convergent thinking as well, because the participants come up with appropriate solutions for which the brick can be used.

3. Characteristics of creativity tasks

As divergent and convergent thinking are in interplay in creativity tasks, we propose a taxonomy that is based on task

characteristics related to these two thinking processes. We distinguish between open-ended and close-ended, and fragmenting and integrating creativity tasks, because these task characteristics, as represented in the four quadrants of the taxonomy, capture fundamental aspects of divergent and convergent thinking on a concrete task level.

3.1. A taxonomy of creativity tasks

For the definition of open- and close-ended creativity tasks we built on the classification of Wakefield (1987, 1989, 1991), who proposed a matrix with four quadrants formed by the combinations of open or closed problems, and open or closed solutions. We defined open- and close-ended tasks in terms of the open- or close-endedness of the solution of the task (or response of the task). That is, the difference between factual problem solving with a correct response (close-ended task), and expressive problem solving in which the problem is solved in one's own way without a correct response (open-ended task) (Wakefield, 1987). In other words, in close-ended tasks, there are one or more predetermined correct responses on a given stimulus, while in open-ended tasks there are no predetermined correct responses on a given stimulus, see Fig. 1.

In the literature, distinctions between open- and close-endednesss are found in relation to creative problem solving. Lin and Shih (2016) defined open-ended and close-ended creative problems in terms of novelty and appropriateness: "an open-ended creativity problem encourages as many and novel ideas as possible … and a close-ended creative problem not only requires novelty but also appropriateness to fulfil the one final correct answer" (p.37). We agree with Lin and Shih that close-ended tasks are often more closely related to the appropriateness than the novelty of the response, and open-ended tasks are often more closely related to the novelty of the response. We argue, however, that both novelty and appropriateness are needed to come up with creative solutions, and thus related to responses on both open- and close-ended tasks (e.g., Cropley, 2006; Guilford, 1967; Runco, 2008; Simonton & Damian, 2013).

Not only do creativity tasks differ as to whether they require open- or close-ended responses, there are also considerable differences between tasks within the open- and close-ended categories. We capture these differences by distinguishing between integrating and fragmenting tasks. Barbot et al. (2016) combined convergent-integrative thinking tasks and divergent-exploratory thinking tasks in measuring creativity. Convergent-integrative tasks ask for combining, integrating or synthesizing of elements in new ways, while divergent-exploratory tasks ask for expanding the range of solutions in creative problem solving (Barbot, Besançon, & Lubart, 2011, 2016; Lubart, Besançon, & Barbot, 2011). We built on this line of thought by differentiating between integrating and fragmenting tasks in our taxonomy.

With these terms, integrating and fragmenting, we focus on the characteristics of the response of the task. Integrating tasks are based on integration, which can be described as 'the composition of elements to a whole' (Urban & Jellen, 1996) or 'coherence of organization' (Lowenfeld, 1960). The shaping, forming, making, or completing of the final creative product is a central element of integrating creativity tasks (Glăveanu, 2013). Furthermore, integrating tasks ask for a single response on different (more than one) task stimuli. The participant must synthesize different elements in a final product. Fragmenting tasks are defined by the participant's

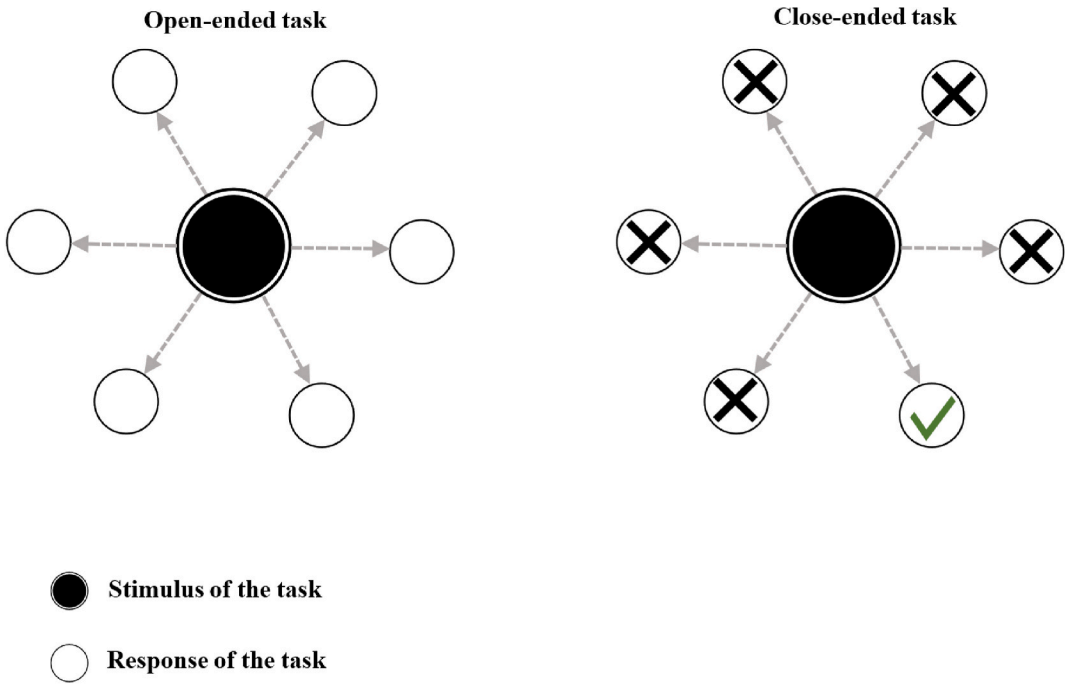


Fig. 1. Visual representations of open-ended and close-ended tasks.

production of different (more than one and often even as much as possible) responses on a single stimulus, see Fig. 2. Fluency (the number of responses) is regularly used as an evaluation-criterium for fragmenting tasks. Furthermore, fragmenting tasks are characterized by the unrelatedness of the responses, as these tasks can be addressed by separate and unrelated reactions.

We combined the axis open- and close-ended tasks with the axis integrating and fragmenting tasks in our taxonomy of creativity tasks, see Table 1. We used the taxonomy to perform a systematic literature review of empirical research on children's creativity development in the primary years.

3.1.1. Examples of creativity tasks of different quadrants of the taxonomy

Quadrant I includes creativity tasks in which there are no predetermined correct responses (open-ended tasks), and the responses combine elements to form a whole (integrating tasks). An example is a storytelling task, in which participants are asked to integrate given elements in their story (EPoC, Lubart et al., 2011).

Quadrant II includes creativity tasks in which there are no predetermined correct responses (open-ended task), and the responses do not have to be related, they consist of separate pieces (fragmenting tasks). An example is the 'Circles and squares task', in which participants are asked to draw as many as possible different drawings with a circle or square as central figure (Guilford, 1967).

Quadrant III includes creativity tasks in which there are predetermined correct responses (close-ended tasks), and the responses combine elements to form a whole (integrating tasks). An example is the 'Progressive Matrices Task', in which participants are asked to finish an incomplete matrix by choosing the right response from six given responses (Progressive Matrices, Raven, Raven, & Court, 1998a; 1998b; 1998c; 1998d).

Quadrant IV includes creativity tasks in which there are predetermined correct responses (close-ended tasks), and the responses do not have to be related, they consist of separate pieces (fragmenting tasks). An example is the 'Creating Equal Numbers task' (CEN), in which participants are asked to make groups with equal numbers with 10 cups in different ways (Tsamir, Tirosh, Tabach, & Leverson, 2009).

3.2. The domain of the task

Besides the task characteristics captured in the taxonomy, we focused on the domain of the task. In previous research, figural tasks have been related to positive developmental patterns, while verbal tasks have been related to slumps (e.g., Said-Metwaly et al., 2020; Zyga et al., 2021). However, concerning the domain of a task, one can distinguish between the domain of the stimulus (what is the domain of the stimulus participants react on?), and the domain of the response (what is the domain of the activity the participants perform?). In their meta-analysis, Said-Metwaly et al. combined the domain of the stimulus and the domain of the response in one variable. In creativity instruments, tasks are sometimes labeled as verbal or figural, while they have figural and verbal elements, for example in the WKCT a task with a verbal response on a figural stimulus is called figural (Wallach & Kogan, 1965). To examine in a more precise manner how the domain of the task relates to development, in the current review we distinguish between the domain of

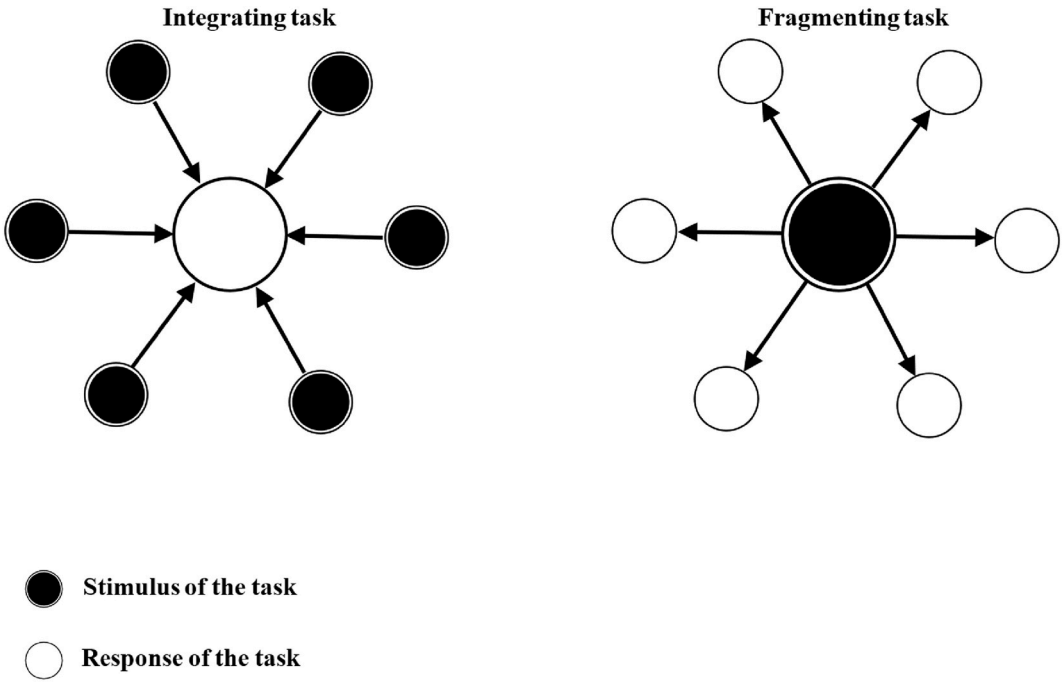


Fig. 2. Visual representations of integrating and fragmenting tasks.

Table 1
A Taxonomy of Creativity tasks.

	Open-ended	Close-ended
Integrating	Quadrant I Creativity tasks in which there is (are) no predetermined correct response(s) (open-ended), and different stimuli ask for a single response that forms a whole (integrating).	Quadrant III Creativity tasks in which there is a (are) predetermined correct response(s) (close-ended), and different stimuli ask for a single response that forms a whole (integrating).
Fragmenting	Quadrant II Creativity tasks in which there is (are) no predetermined correct response (s) (open-ended), and a single stimulus asks for more than one (possibly unrelated) responses (fragmenting).	Quadrant IV Creativity tasks in which there is a (are) predetermined correct response(s) (close-ended), and a single stimulus asks for more than one (possibly unrelated) responses (fragmenting).

the stimulus and the domain of the response of the task.

3.3. The present study

In the present systematic literature review we performed both a descriptive analysis and a meta-analysis on studies of children's creativity development in the primary years, with a focus on task characteristics. We combined these techniques to triangulate findings. The main research questions are:

1. How is the development of children's creativity studied in relation to task characteristics (task taxonomy and task domains) in empirical research of the last 20 years?
2. How can the development of children's creativity be described in relation to task characteristics of the taxonomy and the domain of the task?

4. Method

4.1. Literature search

We followed the Prisma method (Liberati et al., 2009) to conduct the review. A systematic search of the literature was performed in April 2019. We searched in the ERIC, Psycinfo, and Web of Science databases for peer-reviewed journal articles in English, which were published during the last 20 years (1999–2019). We focused on the development of children's creativity within the primary school age range (4–12 years). We used the following key words to search the databases: 'creativity' (such as creativ*, divergent think*, original*, flexib*, fluen*), 'children 4–12 year old' (such as child*, primary education*), and 'development' (such as development*, longitud*, cohort*), see Appendix A for a complete list of search terms. In addition, we added 43 articles from hand-search from leading journals in creativity research (such as *Creativity Research Journal*, *The Journal of Creative Behavior*, and *Thinking Skills and Creativity*) to ensure that relevant articles which were inaccurately indexed were included in the selection. This identification process resulted in 1965 unique articles, see Fig. 3.

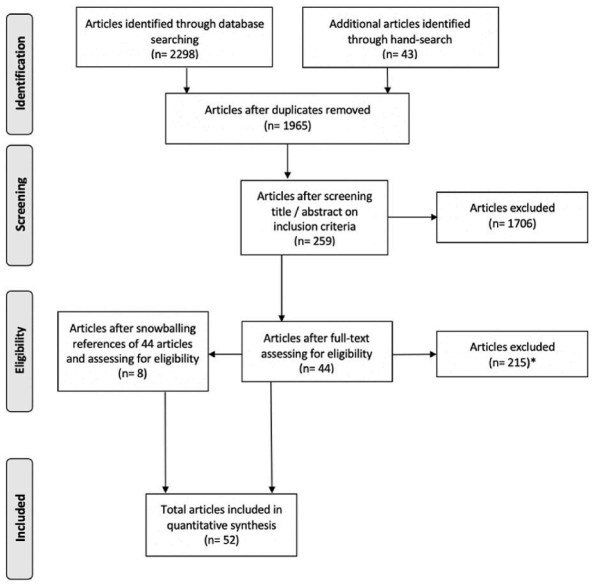


Fig. 3. *Flow diagram of the literature search, screening and inclusion procedure of the articles*
Note. Based on the Prisma method (Liberati et al., 2009) * three articles were not available.

4.1.1. Inclusion criteria

We formulated the following inclusion criteria:

1. The article describes an empirical study in which creativity or divergent thinking is measured.
2. The article describes a study in which 1) creativity is measured in one age cohort for at least two times, these measurements are minimally six months apart, and it is statistically tested whether creativity of the participants differs between these measurements, or 2) creativity is measured in different age cohorts, these age cohorts differ from each other by minimally six months, and it is statistically tested whether creativity differs between these measurements, or it is statistically tested if age or grade has an effect on (or a correlation with) creativity.
3. The article describes a study that contains two measurement-moments within the age range of 4–12 years old, or one measurement-moment within this age range and one (or more) outside the age range. In the latter case, the biggest part (of grades or years) of the investigated age range lies within the 4- to 12-year-old age range.
4. The article describes a study that concerns the regular student population within primary education. Thus, articles describing studies (solely) concerning specific groups of students within primary education, such as students with behavioral problems or gifted students, are not included.
5. When the article describes a study in which developmental patterns of different participant groups are compared, the developmental results concerning the participant group that is most relevant for the current study are included. This means for intervention studies the results of the control group are included, and for studies in which different school systems are compared the results of the school with the most traditional (regular) education are included.

In a first, screening, round, we screened the title and abstract of the articles on the above-described inclusion criteria. If the title and abstract contained too little information for a decision on inclusion, the article was included. This first round rendered 259 articles. In a

Table 2
Coding scheme of instruments and results: overview of variables, categories and codes.

Variables	Categories	Code
Psychometric characteristics instrument	The reliability of the measurement is investigated in the study	
	• Internal consistency reliability of the instrument	cr
	• Interrater reliability of the instrument	irr
	validity of the measurement is investigated in the study	val
Subdimensions of creativity	originality[a]	ori
	fluency	flu
	flexibility	fle
	other subdimensions coded by name	
Type of instrument	questionnaire	quest
	achievement test	achiev
Taxonomy of creativity tasks (see also Table 1)	The majority of tasks is: open-ended integrating	oe-i
	open-ended fragmenting	oe-f
	close-ended integrating	ce-i
	close-ended fragmenting	ce-f
	Equal numbers of tasks are open-ended integrating and open-ended fragmenting	oe-i/f
	oe-f, oe-i, ce-f or ce-i are not applicable	na
Domain of the stimulus of the task: all tasks are coded separately	The core of the stimulus of the task is: figural: includes an imagery stimulus, like a drawing, graphic image, photo, 3d-object and/or video	fig
	verbal: is written or oral	verb
	motor: is a motor situation	mot
	mathematical: is a (numerical) problem	math
	musical: is music with voice and/or instruments	mus
The domain of the response of the task: all tasks are coded separately	The core of the stimulus of the task is: figural: participants make drawings and/or design by drawing	fig
	verbal: participants speak or write including storytelling from imagination	verb
	motor: participants move with their body and/or materials	mot
	mathematical: participants solve a (numerical) problem	math
	musical: participants make music with their voice and/or instruments	mus
	choosing: participants choose a response option	choo
Developmental results: tested developmental pattern of the largest grade range	A tested correlation between creativity and age/grade (negative/no/positive)	grx-y −/=/+
	A tested effect of age/grade on creativity (negative/no/positive)	grx-y −/=/+
	A tested difference between two not consecutive grades (negative/no/positive).	grx-gry −/=/+
	A tested difference between two grade ranges (negative/no/positive)	grc/d-gre/f −/=/+
	(A) tested difference(s) between measurements A (MA) and B (MB) within a grade/year (negative/no/positive)	grx_{MA}-grx_{MB} −/=/+
	(A) tested difference(s) between measurements A (MA) and B (MB) of a grade-range within a grade/year (negative/no/positive)	grx/z_{MA}-grx/z_{MB} −/=/+

Note.
[a] Uniqueness, novelty and originality are much related (Plucker et al., 2004), and therefore the subdimensions novelty and uniqueness are coded as originality.

second round, a full-text screening for eligibility of the remaining articles was performed with the same inclusion criteria. This second round rendered 44 articles. We searched the references of these articles for missing literature (snowballing). After assessing for eligibility with the inclusion criteria, an additional eight articles were included, resulting in 52 articles for analysis. The literature search, screening and inclusion procedure is captured in a diagram, see Fig. 3.

4.2. Coding procedure and scheme

4.2.1. Coding for the descriptive analysis

A coding scheme was developed to systematically analyze the 52 articles in relation to the research questions. The scheme was tested and refined in several rounds until the authors agreed on the (sub)categories and descriptions of the variables. The final scheme consisted of 16 variables. We used eight variables for providing background information on the study: the country, research design (cohort/longitudinal/intervention), number of participants in the total study (n), number of participants within primary education of which the results were included in the analysis (nr; see inclusion criterium 3 and 5), grades of the study's participants, number of cohorts/measurements within primary education, time between measurements in months, and the creativity term mentioned in the article title. The remaining eight variables were used to analyze the used creativity tasks of the instruments, and to map the results on the development of creativity. Table 2 gives an overview of these variables, as well as the categories of the variables, and matching codes that were used in further analyses. These variables are: whether the psychometric characteristics of the used instrument were investigated in the included study, subdimensions that are included in the study's analyses, type of creativity instrument, two variables in relation to the task taxonomy: open-ended versus close-ended, and integrating versus fragmenting (see Table 1), domain of the stimulus of the task, domain of the response of the task, and finally, the study's results concerning creativity and sub-dimensions development. For the results, statistically tested differences concerning the largest grade-range present in the included study are reported, portraying a global picture of the development of creativity. We did so, because we are not interested in a detailed picture of creativity development (specific bumps and slumps), but solely in more global patterns of creativity development in relation to task characteristics.

To the end of determining interrater reliability seven articles (13% of the included articles) were double coded by the first and second author on the variables of the studies. This process rendered percentage agreements of 100% whether the psychometric characteristics of the used instrument were investigated in the included study, 82% for the subdimensions of creativity included in the study's analysis, 91% for type of instrument, 91% for open-ended/close-ended tasks, 91% for integrating/fragmenting tasks, 91% for the domain of the stimulus of the task, 91% for the domain of the response of the task, and 73% for significant developmental results. After establishing the interrater agreement, differences between raters were discussed, until agreed-upon solutions were reached. Finally, all 52 articles were coded on the 16 variables by the first author. In Appendix B, all 52 selected articles are described in terms of the variables from Table 2, and some relevant background variables.

4.2.2. Coding for the meta-analysis

We performed four separate meta-analyses: one for general creativity, originality, fluency and flexibility. The meta-analyses were performed on the included cohort-studies which had sufficient values and statistics to calculate effect sizes. All results from cohort-studies (with sufficient values and statistics), also from intervention-cohort studies and combined cohort and longitudinal studies, are taken into account (n = 32). As we could not distill comparable effect sizes for cohort studies and longitudinal studies (Lipsey & Wilson, 2001), and the number of longitudinal studies was too low for separate meta-analyses, we did not take these studies into account in the meta-analyses.

To calculate standardized mean differences, for each study we coded the creativity means and standard deviations of the lowest and highest grade, and the number of participants of the lowest and highest grade. If means and standard deviations were not available, F-tests or t-tests were used to calculate standardized mean differences (Lipsey & Wilson, 2001). The first and the second author discussed the coded data and resolved any discrepancies to help eliminate errors in coding.

We included five moderators in the meta-analyses: the lowest grade, the grade-range, the domain of the stimulus of the task (figural, verbal, figural/verbal or other), the domain of the response of the task (figural, verbal, figural/verbal or other), and the taxonomy (open-ended fragmenting, open-ended integrating or other). The moderator variables were already coded (see paragraph 4.2.1) and were subsequently used in the meta-analyses.

4.3. Data analysis

To answer the first research question, we conducted a descriptive analysis on instrument characteristics and background variables. We used 'articles' as the unit of analysis. In addition, we used a descriptive analysis on task characteristics (taxonomy and stimulus and response domains). Because articles often report about more than one instrument and the instruments often contain more than one task, we treated these separately when the tasks differed on the for the current review relevant task characteristics (taxonomy and stimulus and response domains), and results were reported separately on these. In relating the study of children's creativity development to task characteristics (see paragraph 5.1.3. and further), we did not include research using questionnaires, as questionnaires could not be categorized on task characteristics (taxonomy and domains).

To answer the second research question, we first performed a descriptive analysis on the reported developmental patterns of general creativity and the creativity subdimensions in the included articles. To this end, we coded statistically tested creativity differences concerning the largest grade-range in each of the included studies. These developmental results are the unit of analysis for the

descriptive analysis of research question two. Next, we related these developmental patterns to the task taxonomy and the domains of the stimulus and response of the tasks. As the included studies varied in whether they reported on general creativity and/or specific subdimensions, as well as in the instruments they applied and the tasks these instruments included (see previous paragraph), and grade-ranges they concerned, most articles rendered multiple results for the descriptive analysis. We included in the analysis, if available, developmental creativity patterns of both general creativity, and the creativity subdimensions originality, fluency, and flexibility. We reported the developmental results per grade-group: kg1-2, gr1-3 and gr4-6. When a result concerned multiple grade groups, we reported the result at all concerning grade-groups.

As the descriptive analysis did not take into account effect sizes of the included studies, a next step in answering the second research question was the performance of meta-analyses on standardized mean differences (n = 84) of general creativity as well as the subdimensions of creativity across grades for the highest grade range. Some of the studies reported multiple standardized mean differences (e.g., for separate subgroups of the data, or for different measurement instruments). The dependency across such effect sizes was taken into account using three-level meta-analysis (Konstantopoulos, 2011; Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013) in R (R Core Team, 2022) using the package metafor (Viechtbauer, 2010). First, we evaluated the average effect size across studies while taking the grade range and starting grade into account by adding those study-level variables as moderators in a meta-regression. We centered grade range around 1 and lowest grade to 1, such that the estimated intercept from the meta-regression would represent studies with a grade range of 1 and a lowest grade being 1. Next, we evaluated the moderating effect of the for the review relevant task characteristics: the task taxonomy ("open-ended integrating", "open-ended fragmenting" versus "other", for general creativity only), and the domain of the stimulus and the domain of the response ("figural", "verbal", "figural/verbal" or "other"). We used a significance level of $\alpha = 0.05$. The complete data and R-syntax and output for the analyses can be found on https://osf.io/t8pwv.

We evaluated possible publication bias by inspecting funnel plot asymmetry. For general creativity and the subdimensions the funnel plots of the effect sizes show a fairly symmetrical distribution around the mean effect, reflecting no clear indications of publication bias, see Appendix C, Fig. C1, C2, C3 and C4. Finally, we applied a meta-regression to test whether results across studies were different on two reliability-related moderators: whether a study used a frequently used creativity instrument or not, and whether a study investigated the reliability of the used creativity instrument or not (for an operational definition of frequently used creativity instruments, and distributions of the reliability-related moderators see paragraph 5.1.2).

5. Results

5.1. How is Children's creativity development studied in relation to task characteristics?

5.1.1. General characteristics

The studies in the included articles have been performed in 19 different countries: the United States (10), Germany (7), Hongkong (6), Poland (5), China (4), Taiwan (3), France (2), Israel (2), Spain (2), The Netherlands (2), and Turkey (2). From the remaining seven countries, one article is included in the review. In 30 articles (58%) the research-design is a cohort study, in seven articles (13%) a combination of a cohort and longitudinal study, in seven articles (13%) a longitudinal study, five articles (10%) are classified as an intervention-longitudinal study, and three articles (6%) as an intervention-cohort study. The number of participants in the articles ranges from 18 to 272,599. Table 3 demonstrates the numbers and percentages of articles reporting on studies with different numbers of participants, grade ranges, and grade groups (kindergarten 1–2, grade 1–3, grade 4–6).

5.1.2. Creativity instruments

In total, 36 different instruments are used in the 52 articles. The Test for Creative Thinking-Drawing Production (TCT-DP; Urban & Jellen, 1996) is used in 12 articles (23%), the Torrance Test of Creative Thinking (TTCT; Torrance, 1966) in nine articles (17%), the Wallach Kochan Creativity Tests (WKCT; Wallach & Kogan, 1965) in six articles (12%). The Creativity Assessment Packet (CAP; Williams, 1980; 1993), the Thinking Creatively in Action and Movement (TCAM; Torrance, 1981), and the Creative Reasoning Task (CRT; Jaarsveld, Lachmann, Hamel, & van Leeuwen, 2010; 2012) are all used in three articles (6%). The remaining 30 instruments (83% of all instruments) are only used once or twice in the articles. Almost all articles (51, 98%) use achievement instruments, and only five (10%) (also) use questionnaires. In 37 articles (71%), (one or more) psychometric characteristics of the used instruments were investigated: in 27 articles (52%) the internal consistency reliability, in 22 articles (42%) the interrater reliability, and in two articles

Table 3
Numbers (and percentages) of articles per participant range, grade range, and grade group.

Participants					Grade range			Grade group		
<50	>50 <100	>100 < 250	>250 < 1000	>1000	0.5–2	3–4	5–7	kg1-2	1–3	4–6
4 (8)	10 (19)	17 (33)	10 (19)	11 (21)	26 (50)	17 (33)	9 (17)	13 (25)	38 (73)	39 (75)

Note. kg1-2 = kindergarten 1–2. Participants: reported total number of participants in the study/studies in the article. Grade range: the largest grade range for which developmental creativity results are tested in the study/studies in the article. Grade groups: combined grades that the study/studies in the article concern(s). Percentages relate to the total number included articles, n = 52. Articles concern frequently multiple grade-groups and thereby the total percentage for grade-groups is higher than 100%.

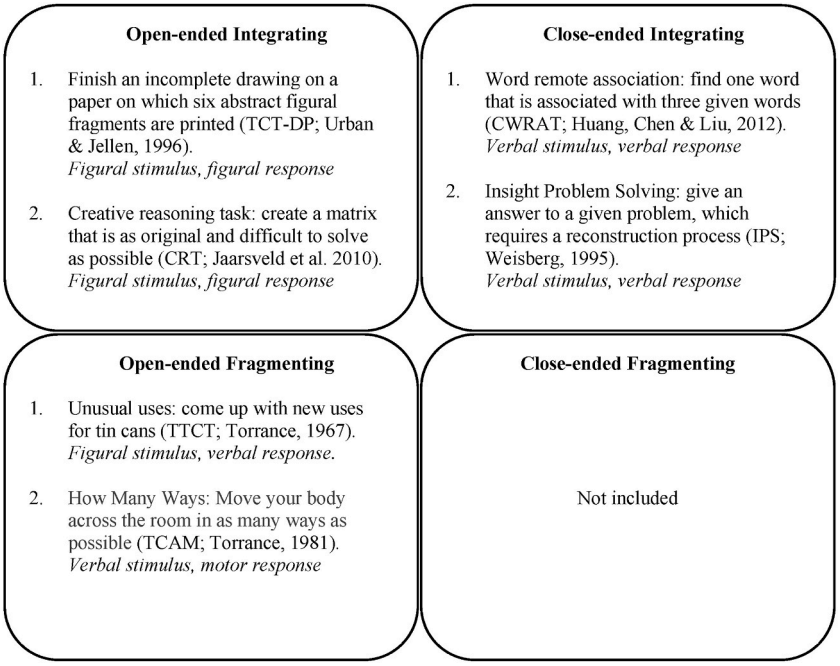


Fig. 4. Examples of the in the included articles used tasks per taxonomy quadrant.
Note. Huang et al., 2012. Weisberg, 1995.

(4%) the validity. For six frequently used creativity instruments (TCT-DP, TTCT, WKCT, CAP, TCAM and CRT), which were each minimally used in three in the review included articles (and in total used in 63% of the articles), the reliability has also been investigated in previous studies (e.g. TCT-DP: Urban, 2004; Urban & Jellen, 1996; TTCT: Torrance et al., 1990; WKCT: Cheun et al., 2004; CAP: Williams, 1993; TCAM: Torrance, 1981; CRT: Jaarsveld, Lachmann, & van Leeuwen, 2012).

Regarding the subdimensions, originality is reported in the analyses in 26 articles (50%), fluency in 25 articles (48%) and flexibility in 19 articles (37%). In 19 articles (37%) another subdimension is reported in the analyses, for example, elaboration or imagination.

5.1.3. Task characteristics: taxonomy and domains of the tasks

In several cases we coded and included more than one task per article in the following analyses (see paragraph 4.3).

In 52 articles (100%) an open-ended creativity task, and in one article (2%) a close-ended creativity task is (also) used. In 34 articles (65%) an open-ended fragmenting task is used, in 23 articles (44%) an open-ended integrating task, in one article (2%) a combination of open-ended integrating and open-ended fragmenting tasks, and in one article (2%) a close-ended integrating task. In Fig. 4 examples of in the included articles used tasks per taxonomy quadrant are described.

Table 4 shows the numbers and percentages of articles including tasks from the different taxonomy categories per grade group. In all grade groups (kindergarten groups, grade groups 1–3 and 4–6) open-ended fragmenting tasks are used most frequently in the articles. Additional, in all grade groups a substantial number of articles included open-ended integrating tasks. Close-ended integrating tasks are only used once in grade groups 1–3 and 4–6, and not in the kindergarten groups. Close-ended fragmenting tasks are not used.

Concerning the domains of the tasks, in 34 articles (65%) a stimulus from the figural domain is used, in 12 articles (23%) a combination of a figural and verbal stimulus, in 13 articles (25%) a verbal stimulus, in three articles (6%) a motor stimulus, in one article (2%) a musical stimulus, and in one article (2%) a mathematical stimulus. In 27 articles (52%) a figural response is used, in 15 articles (29%) a verbal response, in seven articles (13%) a combination of a figural and verbal response, in five articles (10%) a motor response or a response option was chosen, in one (2%) article a musical response, in one (2%) article a mathematical response, and in one (2%) article a

Table 4
Numbers (and percentages) of articles that include tasks from the taxonomy categories per grade-group.

Grade group	Taxonomy categories				
	oe-i	oe-f	oe-i/f	ce-i	ce-f
kg1-2 (n = 15)	6 (40)	8 (53)	1 (7)	- (–)	- (–)
1-3 (n = 44)	19 (43)	23 (52)	1 (2)	1 (2)	- (–)
4-6 (n = 44)	17 (39)	26 (59)	- (–)	1 (2)	- (–)

Note. N = 52 articles. Articles can concern multiple grade-groups and can contain multiple tasks from different taxonomy categories, resulting in higher totals compared to Table 3. Kg1-2 = kindergarten 1–2, oe-i = open-ended integrating, oe-f = open-ended fragmenting, oe-i/f = open-ended integrating and fragmenting, ce-i = close-ended integrating, ce-f = close-ended fragmenting.

Table 5
Numbers (and percentages) of articles reporting on general creativity and subdimensions of creativity per open-ended taxonomy quadrants and stimulus and response domain of the task.

Stimulus domains of the task							Response domains of the tasks							
	fig	verb	mot	mus	math	fig/ verb		Fig	verb	mot	mus	math	fig/ verb	fig/ verb/ mot
General creativity oe-i (n = 23)	19 (83)	2 (9)	–	–	–	2 (9)	General creativity oe-i (n = 24)	19 (80)	3 (13)	–	–	–	2 (8)	–
General creativity oe-f (n = 40)	17 (43)	8 (20)	3 (8)	1 (3)	1 (3)	10 (25)	General creativity oe-f (n = 38)	10 (26)	15 (39)	5 (13)	1 (3)	1 (3)	5 (13)	1 (3)
Originality oe-f (n = 25)	15 (60)	4 (16)	1 (4)	–	–	5 (20)	Originality oe-f (n = 25)	10 (40)	9 (36)	3 (12)	–	–	3 (12)	–
Fluency oe-f (n = 33)	16 (48)	8 (24)	1 (3)	–	1 (3)	7 (21)	Fluency oe-f (n = 31)	10 (32)	13 (42)	3 (10)	–	1 (3)	3 (10)	1 (3)
Flexibility oe-f (n = 22)	12 (55)	5 (23)	1 (5)	–	–	4 (18)	Flexibility oe-f (n = 21)	7 (33)	9 (43)	1 (5)	–	–	3 (14)	1 (5)

Note. oe-i = open-ended integrating, oe-f = open-ended fragmenting, fig = figural, verb = verbal, mot = motor, mus = musical, math = mathematical, fig/verb = figural and verbal, fig/verb/mot = figural and verbal and motor.

Table 6
Development of general creativity per grade-group: numbers (and percentages) of results.

Grade group	Negative development	No development	Positive development	Total[a] (n = 83)
kg1-2	–	3 (33)	6 (67)	9
1–3	1 (3)	11 (29)	26 (68)	38
4–6	1 (3)	9 (25)	26 (72)	36

Note. kg1-2 = kindergarten 1–2.
[a] This total includes results of the close-ended integrating (ce-i) and open-ended mixed integrating and fragmenting (oe-i/f) taxonomy categories: one result in kg1-2, three results in gr1-3, and two results in gr4-6.

combination of figural, verbal and motor response.

Table 5 shows the number and percentage of articles reporting general creativity and subdimensions of creativity per taxonomy quadrant and stimulus and response domain of the task. In this Table, we focus on the two open-ended taxonomy-quadrants, as the number of articles reporting results within the close-ended and mixed open-ended fragmenting and integrating taxonomy categories was limited. For measuring general creativity with open-ended integrating tasks, most frequently figural stimuli, and figural responses are used. For measuring general creativity with open-ended fragmenting tasks, most frequently figural stimuli, but verbal responses are used.

For measuring the subdimensions originality, fluency and flexibility, only the results in the open-ended fragmenting taxonomy quadrant are presented (97% of the subdimensions is mapped as open-ended fragmenting). For measuring all subdimensions most frequently figural stimuli are used. For measuring originality most frequently figural responses are used, while these are verbal responses for measuring fluency and flexibility.

5.2. How can creativity development be described in relation to task characteristics?

5.2.1. Developmental patterns of general creativity

Descriptive Analysis of Developmental Patterns of General Creativity. Table 6 shows, for each grade-group, the numbers and percentages of general creativity results demonstrating the different developmental patterns (negative, no, positive). Across all grade-groups, the majority of the results of general creativity demonstrate positive development. However, it should be noted that a substantial number of results in all grade-groups shows no development of general creativity.

Meta-analysis of Developmental Patterns of General Creativity. Additionally, we conducted a meta-analysis on standardized mean differences of general creativity across grades for the largest grade range, on the subset of the cohort studies. There were 50 effect

Table 7
Development of creativity subdimensions per grade-group: numbers (and percentages) of results.

Grade group	originality (oe-f) (n = 42)			fluency (oe-f) (n = 61)			flexibility (oe-f) (n = 37)		
	neg	no	pos	neg	no	pos	neg	no	pos
kg1-2	- (–)	4 (67)	2 (33)	- (–)	4 (40)	6 (60)	- (–)	1 (25)	3 (75)
1–3	1 (7)	9 (60)	5 (33)	2 (9)	8 (35)	13 (57)	2 (17)	3 (25)	7 (58)
4–6	3 (14)	11 (52)	7 (33)	4 (14)	9 (32)	15 (54)	2 (10)	8 (38)	11 (52)

Note. kg1-2 = kindergarten 1–2, oe-f = open-ended fragmenting, neg = negative development, no = no development, pos = positive development.

Table 8
Development of general creativity per grade-group and per open-ended taxonomy quadrant: numbers (and percentages) of results.

Grade group	oe-i general creativity(n = 48)			oe-f general creativity (n = 29)		
	neg	no	pos	neg	no	pos
kg1-2	- (–)	1 (17)	5 (83)	- (–)	2 (100)	- (–)
1–3	- (–)	7 (32)	15 (68)	1 (8)	4 (31)	8 (62)
4–6	- (–)	7 (35)	13 (65)	1 (7)	2 (14)	11 (79)

Note. kg1-2 = kindergarten 1–2, oe-i = open-ended integrating, oe-f = open-ended fragmenting, neg = negative development, no = no development, pos = positive development. Results with close-ended and mixed open-ended fragmenting and integrating categories are not reported in this Table.

sizes obtained from 22 independent studies. The observed standardized mean differences ranged from −1.07 to 1.56, with the majority of estimates being positive (86%). The three-level meta-analysis showed that there was significant heterogeneity across effect sizes, with a between study-variance of $\tau^2 = 0.12$, and no significant variance at the effect size level ($\tau^2 = 0.00$). The moderators 'grade range' and 'lowest grade' (see method paragraph 4.3) did not have a statistically significant effect ($\chi^2(2) = 1.87$, $p = .39$). The average standardized mean difference pertaining to a grade range of 1 and lowest grade of 1 was positive and statistically significant, $\widehat{\mu} = 0.59$ (SE = 0.13, $p < .05$, 95% CI = [0.34; 0.85]). We provided a three-level forest plot of standardized mean differences of general creativity in Appendix D, Fig. D1.

To summarize, both the descriptive and the meta-analysis demonstrate a positive developmental pattern of general creativity across primary education.

5.2.2. Developmental patterns in relation to task characteristics

Descriptive Analysis of Developmental Patterns of Creativity Subdimensions (open-ended fragmenting taxonomy quadrant). Table 7 shows, for each grade-group, the numbers and percentages of results of subdimensions of creativity demonstrating the different developmental patterns (negative, no, positive). Importantly, the reported results of the subdimensions originality, fluency, and flexibility (97%) where all categorized in a single taxonomy quadrant: the open-ended fragmenting quadrant. The majority of the results concerning the originality subdimension show no development, and this pattern is found in all grade-groups. The majority of results concerning the fluency and flexibility subdimensions show positive development in all grade groups.

Meta-Analyses of Developmental Patterns of the Subdimensions of Creativity (open-ended fragmenting taxonomy quadrant). We conducted three meta-analyses on standardized mean differences of the subdimensions of creativity across grades for the largest grade range, on the subset of the cohort studies. For originality we obtained 24 observed effect sizes from 15 independent studies. The observed standardized mean differences ranged from −0.28 to 0.90, with the majority of estimates being positive (79%). The three-level meta-analysis showed that the omnibus test for moderation was not significant ($\chi^2(2) = 5.10$, $p = .08$), but the univariate test of lowest grade showed that the average effect sizes for lower grades were significantly higher than the effect sizes for higher grades ($\widehat{\beta} = -0.08$, $p = .03$). The heterogeneity of effect sizes was larger at the effect size level ($\tau^2 = 0.04$) than at the study level ($\tau^2 = 0.00$). The average standardized mean difference pertaining to a grade range of 1 and lowest grade of 1 was positive and statistically significant ($\widehat{\mu} = 0.42$ (SE = 0.09, $p < .05$, 95% CI = [0.24; 0.61]).

For fluency we obtained 33 effect sizes representing standardized mean differences in fluency, obtained from 15 studies. These effect sizes ranged from −0.47 to 1.26, with 91% of them being positive. The moderators 'grade range' and 'lowest grade' did not have significant effects on the effect sizes ($\chi^2(2) = 2.97$, $p = .23$). The estimated variance at the effect size level was $\tau^2 = 0.075$, while it was estimated at zero at the study level. The average standardized mean difference pertaining to a grade range of 1 and lowest grade of 1 was positive and statistically significant $\widehat{\mu} = 0.41$ (SE = 0.08, $p < .05$, 95% CI = [0.26; 0.56]).

For flexibility the 15 observed effect sizes of flexibility came from 11 independent studies. The standardized mean differences ranged from −0.64 to 1.22, and most of them were positive (87%). The three-level meta-analysis showed that there was significant heterogeneity, with a between effect size variance of $\tau^2 = 0.16$, and no significant variance at the study level ($\tau^2 = 0.00$). The moderators 'grade range' and 'lowest grade' did not have a statistically significant effect ($\chi^2(2) = 1.99$, $p = .37$). The average standardized mean difference pertaining to a grade range of 1 and lowest grade of 1 was positive and statistically significant $\widehat{\mu} = 1.05$ (SE = 0.49, $p < .05$, 95% CI = [0.10; 2.00]). We provided three-level forest plots of standardized mean differences of the subdimensions in Appendix D, Fig. D2, D3 and D4.

To summarize, both the descriptive and the meta-analyses demonstrate a somewhat different developmental pattern for originality as compared to fluency and flexibility.

Descriptive Analysis of Developmental Patterns of General Creativity Related to the Taxonomy. Table 8 shows, for each grade-group, the numbers and percentages of results of general creativity demonstrating the different developmental patterns (negative, no, positive) per taxonomy quadrant. We focus here on each of the two open-ended taxonomy-quadrants, as the number of results within the close-ended and mixed open-ended fragmenting and integrating taxonomy categories was limited (total of six results). The majority of results of general creativity measured with open-ended integrating tasks demonstrate positive development, and this pattern is found across all grade-groups. For general creativity measured with open-ended fragmenting tasks a similar pattern is found, with the exception of the kindergarten grade group, but the total number of results is very small here. Furthermore, a substantial number of results of general creativity measured with open-ended integrating and fragmenting tasks shows no development across grade-groups.

Descriptive Analysis of Developmental Patterns of General Creativity Related to Task Domains. Table 9 shows, for each

Table 9
Development of general creativity per grade group and per stimulus and response domains of the task: numbers (and percentages) of results.

	General creativity Stimulus of the task									General creativity Response of the task								
	fig*			verb*			fig/verb*			fig*			verb*			fig/verb*		
	neg	no	pos	neg	no	pos	neg	no	pos	neg	no	pos	neg	no	pos	neg	no	pos
Kg1-2	-	3	6	-	-	-	-	-	-	-	1	5	-	1	-	-	1	1
	(–)	(33)	(67)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(17)	(83)	(–)	(50)	(–)	(–)	(50)	(50)
Gr1-3	-	5	17	1	2	2	-	2	4	-	5	14	1	1	6	-	3	3
	(–)	(23)	(77)	(20)	(40)	(40)	(–)	(33)	(67)	(–)	(26)	(74)	(13)	(13)	(75)	(–)	(50)	(50)
Gr4-6	-	6	14	1	1	2	-	2	7	-	5	12	1	1	8	-	3	3
	(–)	(30)	(70)	(25)	(25)	(50)	(–)	(22)	(78)	(–)	(29)	(71)	(10)	(10)	(80)	(–)	(50)	(50)

Note. fig = figural, verb = verbal, fig/verb = figural and verbal, neg = negative development, no = no development, pos = positive development. * From the results with close-ended tasks four results had a verbal stimulus and a verbal response. From the mixed fragmenting and integrating category two results had a figural stimulus and a mixed figural and verbal response.

grade-group, the numbers and percentages of results of general creativity demonstrating the different developmental patterns (negative, no, positive) per stimulus and response domain of the task. We focus on figural, verbal and mixed figural and verbal domains of the task (stimulus and response) here, as the number of results concerning other domains were limited (see Table 5). Most of these results show positive development, which is in line with the results of developmental patterns of general creativity (see Table 6), with the exception of most of the results of general creativity with a verbal stimulus of the task in grade 1–3 and the results of general creativity with a figural/verbal response of the task in all grade groups. In these cases, equally often positive development and no development is found.

Meta-Analysis with Task Characteristics as Moderators. The effect of three task variables (taxonomy, the domain of the stimulus of the task, and the domain of the response of the task) was examined on the developmental differences in general creativity and subdimensions of creativity in four separate moderator analyses. There were no significant moderating effects on the general creativity outcomes ($\chi^2(10) = 9.92$, $p = .45$), and on the originality outcomes ($\chi^2(10) = 9.93$, $p = .45$). The estimated effect of open-ended integrating tasks was positive, but not significant. For fluency, the omnibus test for moderation was not significant ($\chi^2(8) = 11.43$, $p = .18$), but the univariate test of the type of response showed that the 10 effect sizes with type of response "figural/verbal" were on average significantly lower than the 10 effect sizes with type of response "figural", which served as the reference group ($\widehat{\beta} = -0.83$, $p = .03$), after controlling for the effects of the other moderators. A similar result was found for the subdimension flexibility, where the omnibus test for moderation was also not significant ($\chi^2(8) = 8.83$, $p = .36$), but the three effect sizes with type of response "figural/verbal" were on average significantly lower than the three effect sizes with type of response "figural" ($\widehat{\beta} = -1.40$, $p = .02$), after controlling for the effects of the other moderators. There was no statistically significant moderation of the two reliability-related moderators on general creativity or the three subdimensions.

To summarize, in both the descriptive and the meta-analyses the use of a combination of a figural and verbal response of the task is related to a less pronounced developmental pattern compared to when a figural response is used. This picture emerged for general creativity in the descriptive analysis and for the subdimensions fluency and flexibility in the meta-analyses.

6. Conclusion and discussion

By answering the first research question we provided an overview of the empirical research of the last 20 years on children's creativity development in relation to task characteristics. The body of research (52 articles) demonstrated a varied and fragmented picture, especially concerning the measurement of creativity and its subdimensions. A vast array of substantially different creativity instruments (36) has been used in the articles, with six instruments being used in about two third of the articles: the TCT-DP (Urban & Jellen, 1996), TTCT (Torrance, 1966), WKCT (Wallach & Kogan, 1965), CAP (Williams, 1993); TCAM (Torrance, 1981) and CRT (Jaarsveld et al., 2012). About two-thirds of the included articles concerned one or more of the examined subdimensions of creativity: originality, fluency and flexibility. Additionally, in approximately a third of the articles other subdimensions have been included, such as elaboration or imagination. There are substantial differences between the studies in sample sizes, included grades, and grade-ranges. For example, a quarter of the articles concerned kindergarten, whereas the grade-groups 1–3 and 4–6 are both included in three quarter of the articles.

Given the large variety of instruments used in the reviewed studies, often consisting of several tasks, the concrete task-centered taxonomy that was developed for the purpose of this review was helpful to provide more clarity in distinguishing between different types of creativity tasks. We distinguish between open-ended and close-ended, and fragmenting and integrating creativity tasks, because these task characteristics capture fundamental aspects of divergent and convergent thinking on a concrete task level. The taxonomy is thereby in line with a broad conceptualization of creativity, not only including divergent but also convergent thinking (e.g., Barbot, 2019; Runco & Acar, 2012). Consequently, in contrast to reviews purely focusing on divergent thinking tasks (e.g., Said-Metwaly et al., 2020), the current task-centered review focuses on a broad range of creativity tasks, also including open-ended integrating tasks, such as the TCT-DP.

The results of the current review study demonstrated that tasks categorized in the open-ended integrating and the open-ended fragmenting quadrants of the taxonomy were both applied frequently in the studies. Close-ended tasks (both integrating and fragmenting tasks) were less applied. This could be related to the fact that close-ended tasks ask for problem solving with a correct response, and this process is less closely associated with the creativity criterium of novelty (Wakefield, 1987). Close-ended tasks that are included in the taxonomy concern tasks where the right response(s) need(s) an insight, by for example reframing the problem, or coming up with a (new) idea through association (Wu & Cheng, 2019). This reframing of a problem, or so-called problem finding, has been related to creativity (Abdulla, Paek, Cramond, & Runco, 2018; Wakefield, 1987), and one could argue that in reframing of a problem an element of novelty is included. We see that in all grade groups mostly open-ended fragmenting tasks were used. But also open-ended integrating tasks are frequently applied in all grade groups.

Besides analyzing tasks by means of the taxonomy, in the current review we also analyzed tasks in terms of the used domains of the stimulus and response of the tasks. In the majority of the included articles, figural stimuli and responses have been used. Zooming in on general creativity measured with open-ended integrating tasks, our results showed that these tasks also mostly included figural stimuli and responses. For measuring general creativity with open-ended fragmenting tasks, the stimuli were mostly figural, but the responses mostly verbal.

In answering the second research question concerning developmental patterns of creativity and subdimensions in relation to task characteristics (task taxonomy and domains of the tasks), we performed a descriptive analysis (on all included articles) and a meta-analysis (on a substantial subset of the included articles). In this way we were able to triangulate our findings. In line with

previous research (e.g., Besançon, Lubart, & Barbot, 2013; Lau & Cheung, 2010a, 2010b; Maker et al., 2008; Said-Metwaly et al., 2020), we found that children's general creativity increased across primary education, including kindergarten. Similarly, we saw an increase of general creativity with age when creativity was measured with open-ended integrating and with open-ended fragmenting tasks. An exception to this developmental pattern was found for the subdimension originality in the open-ended fragmenting taxonomy quadrant. The descriptive analysis demonstrated that the majority of originality results in all grade groups showed no development, which contrasted with the majority of fluency and flexibility results demonstrating positive development. In line with that finding, the meta-analysis for originality demonstrated that the majority of the (average) effect sizes per study were not significantly larger than zero, while for fluency and flexibility most study-level effect sizes were significantly larger than zero. Additionally, in the originality meta-analysis the lowest grade was found to be a significant moderator in the univariate analysis. That is, the lower the grade, the higher the average effect sizes. Said-Metwaly et al. (2020) also included the creativity subdimensions in their meta-analysis, and reported similar developmental patterns for the creativity subdimensions, including originality, and overall divergent thinking. However, a close inspection of the results of Said-Metwaly et al. also suggests less pronounced development for originality as compared to fluency and flexibility.

An explanation for the less pronounced developmental pattern in originality might be found in Piaget's developmental theory (1951), and especially in the existence of the concrete operational stage in middle childhood. It is hypothesized that during this concrete operational stage children's creativity does not increase, because of the attention to and the emergence of logical and concrete thinking (e.g. Lubart & Lautrey, 1995; Rosenblatt & Winner, 1988). The fact that we found less pronounced development for originality as compared to general creativity and the other subdimensions could possibly be explained by originality being very closely related to the novelty creativity criterium (Plucker et al., 2004); without originality one cannot be creative (Nijstad, De Dreu, Rietzschel, & Baas, 2010).

Looking at the domains of the tasks, the descriptive analyses demonstrated that most general creativity results for all stimuli and response domains showed positive development, with the exception of the results with a verbal stimulus in grade 1–3 and the results with a figural/verbal response in all grade groups equally often showing positive and no development. In line with these findings, the meta-analyses for the fluency and flexibility subdimensions demonstrated that the effect sizes concerning combined figural and verbal responses were lower than the effect sizes concerning figural responses. These findings are in line with previous studies demonstrating slumps in creativity development being primarily associated with the use of verbal tasks (Said-Metwaly et al., 2020; Zyga et al., 2021). Our study adds to this literature by distinguishing between the domain of the stimulus and the domain of the response of the task and finds that the less pronounced development when verbal or a combination of figural and verbal tasks are used, is mostly related to the response domains of the task.

6.1. Limitations

A first limitation of the current study is related to the dichotomous character of the taxonomy: the results were categorized in quadrants instead of mapped on a more continuous scale representing degrees of open/close-endedness and integrating/fragmenting. This choice can possibly explain the few tasks being categorized as close-ended, because only when a task asked for predetermined correct responses, and not when it was more geared towards a correct response, it was coded as close-ended. Although the dichotomous nature of the taxonomy was helpful to relate developmental outcomes to task characteristics in the current review, by using the taxonomy in its current form we lost some nuance with regards to the open and close-endedness of tasks.

Second, several effects of the meta-analyses were only shown as univariate effects, while the omnibus tests were nonsignificant. When in future work the studies included in meta-analyses on creativity development in relation to task characteristics will be updated and expanded, these future analyses will give more clarity on the robustness of these univariate effects.

Third, we expected that the effect sizes would have been moderated by grade range (Lau & Cheung, 2010a, 2010b). The most likely cause of not finding such an effect is a lack of statistical power, which is quite common for moderator analyses (Hedges & Pigott, 2004). In our case, the power may be low because of a relatively low number of studies, because of limited variation in the grade range, or because the effect sizes may not be linearly related to the grade range. Indications for a non-linear development of divergent thinking are found (Torrance, 1968). Still, adding these moderator variables allowed us to interpret the intercept as the average effect sizes in studies with the same grade range and same lowest grade.

6.2. Implications for educational practice

The taxonomy is not only useful for research purposes, but also provides a framework that can be used to develop creativity activities for educational practice. If teachers are aware of the spectrum of task characteristics related to creativity, this can help them in the design of authentic creativity activities in the classroom. First, the distinction between open- and close-ended tasks can be helpful in analyzing and discussing the current and desired situation in relation to the use of creativity-provoking activities in the classroom. It is, for example, likely that most activities children in primary education engage in are close-ended in nature, but these are not creativity-provoking activities. This knowledge allows for opportunities for designing a broad spectrum of creativity activities by exploring whether more open-ended activities can be included in the classroom, and whether some of the existing close-ended activities can be transformed into close-ended creativity tasks. Second, creativity is easily associated with the generation of many unrelated ideas (open-ended fragmenting tasks), like in brainstorms (Smith, 1998). However, synthesizing information into one meaningful product (open-ended integrating tasks), can also require creativity. Further awareness of the spectrum of task characteristics related to creativity, might lead teachers to explore how integrating creativity activities can become more prominent in their

teaching. Integrating creativity activities can be included in meaningful education based on authentic real-life problems, which is receiving more and more attention in recent decades (e.g., Francom, 2017; Merrill, 2002, 2007; Van Merrienboer & Kirschner, 2007).

As creativity is increasingly seen as one of the core skills for children to develop in primary education (e.g., Beghetto, 2016; Saywer, 2012), this systematic review provides relevant knowledge about developmental patterns of creativity in relation to task characteristics. Moreover, this review rendered a taxonomy of creativity tasks, which is both useful for clarifying discussions in the creativity research discourse, and the design of creativity activities for the educational practice.

Authorship contribution statement

Olga T.A. Potters: Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing, Tessa J.P van Schijndel: Conceptualization, Methodology, Formal analysis, Writing - review & editing, Suzanne Jak: Methodology, Formal analysis, Writing - review & editing, Joke Voogt: Conceptualization, Methodology, Writing - review & editing.

Funding

This work was supported by ArtEZ University of the Arts, the Dutch Research Council (NWO) project number 023.012.034, awarded to Olga T.A. Potters and project number VI.Vidi.201.009, awarded to Suzanne Jak.

Declaration of competing interest

None.

Data availability

The data supporting the meta-analyses are available at https://osf.io/t8pwv

Acknowledgments

We would like to thank Francesca Casertano for her assistance with the coding process for the meta-analysis.

Appendix A

Search Terms Used in Different Databases

ERIC

Ovid, 1965 to January 2019.

#1 creativity. creativity/OR creative thinking/OR creative development/OR (creativ* OR divergent think* OR ((original* OR flexib* OR fluen*) ADJ3 thinking)).ti,ab,id.

#2 children (4–12). (preschool education OR grade 1 OR grade 2 OR grade 3 OR grade 4 OR grade 5 OR grade 6 OR elementary education OR middle schools).el. OR preschool children/OR young children/OR children/OR (preschool* OR kindergarten* OR child* OR kid OR kids OR teen* OR girl* OR boy* OR ((elementary OR primary education OR primary school*) AND (student* OR participant*)) OR 1st-grade* OR first-grade* OR grade 1 OR grade one OR 2nd-grade* OR second-grade* OR grade 2 OR grade two OR 3rd-grade* OR third-grade* OR grade 3 OR grade three OR 4th-grade* OR fourth-grade* OR grade 4 OR grade four OR 5th-grade* OR fifth-grade* OR grade 5 OR grade five OR 6th-grade* OR sixth-grade* OR grade 6 OR grade six).ti,ab,id.

#3 development. Development*/OR child development/OR developmental stages/OR individual development/OR creative development/OR longitudinal studies/OR followup studies/OR cohort analysis/OR (longitud* OR cohort* OR followup OR follow-up OR (development* ADJ3 (stage* OR creativ*)) OR (trend* ADJ3 (development* OR creativ*)) OR child development OR individual development).ti,ab,id.

PsycINFO

Ovid, 1806 to April Week 1 2019.

#1 creativity. Creativity/OR divergent thinking/OR (creativ* OR divergent think* OR ((original* OR flexib* OR fluen*) ADJ3 thinking)).ti,ab,id.

#2 children (4–12). (school age 6 12 yrs).ag. OR (preschool* OR kindergarten* OR child* OR kid OR kids OR teen* OR girl* OR boy*

OR ((elementary OR primary education OR primary school*) AND (student* OR participant*)) OR 1st-grade* OR first-grade* OR grade 1 OR grade one OR 2nd-grade* OR second-grade* OR grade 2 OR grade two OR 3rd-grade* OR third-grade* OR grade 3 OR grade three OR 4th-grade* OR fourth-grade* OR grade 4 OR grade four OR 5th-grade* OR fifth-grade* OR grade 5 OR grade five OR 6th-grade* OR sixth-grade* OR grade 6 OR grade six).ti,ab,id.

#3 development. Development*/OR childhood development/OR early childhood development/OR developmental stages/OR developmental age groups/OR longitudinal studies/OR followup studies/OR cohort analysis/OR (longitud* OR cohort* OR followup OR follow-up OR (development* ADJ3 (stage* OR creativ*)) OR child development OR individual development).ti,ab,id.

Web of Science

Thomson Reuters, Web of Science Core Collection.

#1 creativity. TS=("creativ*" OR "divergent think*" OR (("original*" OR "flexib*" OR "fluen*") NEAR/2 "thinking"))

#2 children (4–12). TS=("preschool*" OR "kindergarten*" OR "child*" OR "kid" OR "kids" OR "teen*" OR "girl*" OR "boy*" OR (("elementary" OR "primary education" OR "primary school*") AND ("student*" OR "participant*")) OR ″1st-grade*" OR "first-grade*" OR "grade 1″ OR "grade one" OR ″2nd-grade*" OR "second-grade*" OR "grade 2″ OR "grade two" OR ″3rd-grade*" OR "third-grade*" OR "grade 3″ OR "grade three" OR ″4th-grade*" OR "fourth-grade*" OR "grade 4″ OR "grade four" OR ″5th-grade*" OR "fifth-grade*" OR "grade 5″ OR "grade five" OR ″6th-grade*" OR "sixth-grade*" OR "grade 6″ OR "grade six")

#3 development. TS=("longitud*" OR "cohort*" OR "followup" OR "follow-up" OR ("development*" NEAR/2 ("stage*" OR "creativ*")) OR "child development" OR "individual development")

Appendix B

The Selected Articles and a Summary of Core Variables

Nr	Author(s) (year) Country	Design	Measurements (time between*)	Participants	Abbreviation instrument**	Psychometric characteristics	Type of instrument	Domain stimulus	Domain response	Task taxonomy	Developmental results
1	Alacapinar (2013) Turkey	coh	4	n: 172 nr: 112 gr3, 4, 5, 6	TTCT		achiev	fig	verb fig	oe-f	gen: gr3-gr6= ori: gr3-gr6= flu: gr3-gr6= fle: gr3-gr6−
2	Alfonso-Benlliure and Santos (2016) Spain	coh	6	n: 1491 gr1, 2, 3, 4, 5, 6	TCI	irr	achiev	fig	fig	oe-i	gen: gr1-6+ ori: gr1-6+
3	Besançon and Lubart (2008) France	coh long	coh: 4 long: 2 (12)	n: 211 gr1, 2, 3, 4, 5	TTCT (part)	cr	achiev	fig	verb fig	oe-f	coh gen: gr1-5+
					ITT	irr	achiev	verb fig	verb fig	oe-i	coh gen: gr1-5+
4	Chae (2003) Korea, South	coh	3	n: 1366 kg1, 2, gr1	TCT-DP		achiev	fig	fig	oe-i	gen: kg1-(gr)1+
5	Chan et al. (2001) Hong Kong	coh	3	n: 462 gr1, 2, 3	WKCT (part)	cr	achiev	verb	verb	oe-f	flu: gr1-3+
		coh			WKCT (part)	cr	achiev	fig	verb	oe-f	flu: gr1-3=
6	Chan and Zhao (2010) Hong Kong	coh	2	n: 223 nr: 164 gr1/5, 6/11	CDAT (part)	irr	achiev	verb	fig	oe-i	gen: gr1-5=
					CCRS		quest	verb	choo	na	gen: gr1-5=
					WKCT (part)		achiev	verb fig	verb	oe-f	gen: gr1-5+
7	Chang, Chen, Wu, Chang, and Wu (2017) Taiwan	coh	2	n: 1674 nr: 357 gr5, 6	NTCT	irr	achiev	fig	fig	oe-f	ori: gr5-gr6= flu: gr5-gr6= fle: gr5-gr6=
					CAP (part)		quest	verb	choo	na	gen:gr5-gr6−
8	Charles and Runco (2001) United States	coh	3	n: 102 gr3, 4, 5	WKCT (part)	irr	achiev	verb	verb	oe-f	ori: gr3-5+ flu: gr3-5+ fle: gr3-5+
					ET(ori)		quest	verb	choo	na	gen: gr3-5+
					ET(pref)		quest	verb	choo	na	No results applicable

(*continued on next page*)

(*continued*)

Nr	Author(s) (year) Country	De-sign	Measurements (time between*)	Participants	Abbreviation instrument**	Psychometric characteristics	Type of instrument	Domain stimulus	Domain response	Task taxonomy	Developmental results
9	Claxton, Pannells, and Rhoads (2005) United States	interv-coh	3 (24,60)	n: 184 nr: 124 gr4, 6	CAP (part)		achiev	fig	fig	oe-f	gen: gr4-gr6= ori: gr4-gr6-flu: gr4-gr6= fle: gr4-gr6=
					CAP (part)		quest	verb	choo	na	gen: gr4-gr6=
10	Domíguez, Díaz-Pereira, and Martínez-Vidal (2015) Spain	coh	3	n: 84 gr1/2, 3/4, 5/6	TCAM/BTMT		achiev	verb fig	mot	oe-f	gen: gr1/2-gr5/6+ ori: gr1/2-gr5/6= flu: gr1/2-gr5/6+ fle: gr1/2-gr5/6+
11	Duh and Budefeld (2018) Slovenia	interv-coh	2 (8)	n: 18 gr2	KRDC		achiev	verb	fig	oe-i	gen: $gr2_{MA}$-$gr2_{MB}$=
12	Dziedziewicz, Gajda, and Karwowski (2014) Poland	interv-coh	2 (9)	n: 122 nr: 56 gr3/6	TTCT (part)	cr irr	achiev	fig	fig	oe-f	ori: $gr3/6_{MA}$-$gr3/6_{MB}$= flu: $gr3/6_{MA}$-$gr3/6_{MB}$= fle: $gr3/6_{MA}$-$gr3/6_{MB}$=
					FDCT	cr	achiev	fig	fig	oe-i	gen: $gr3/6_{MA}$-$gr3/6_{MB}$=
13	Dziedziewicz, Oledzka, and Karwowski (2013) Poland	interv-coh	3	n: 59 kg1, 2, gr1	TTCT (part)	irr	achiev	fig	fig	oe-f	ori: kg1-(gr)1= flu: kg1-(gr)1= fle: kg1-(gr)1=
					FDCT	irr	achiev	fig	fig	oe-i	gen: kg1-(gr)1=
14	Gralewski et al. (2016) Poland	coh	8	n: 4854 nr: 2477 kg1, 2, gr1, 2, 3, 4, 5, 6	TCT-DP	cr	achiev	fig	fig	oe-i	gen: kg1-(gr)6+
15	Gubbels, Segers, and Verhoeven (2017) Netherlands	interv-coh	3 (10,22)	n: 116 gr5, 6	ATB (part)	cr	achiev	verb fig	verb	oe-f	gen: gr5-gr6+
16	Hamlen (2009) United States	coh	2	n: 118 gr4, 5	TTCT (part)		achiev	fig	verb	oe-f	ori: gr4-gr5= flu: gr4-gr5= fle: gr4-gr5=
					TTCT (part)		achiev	fig	fig	oe-f	ori: gr4-gr5− flu: gr4-gr5−
17	He (2018) Hong Kong	coh long	long: 4 (12,24,36)	n: 985 nr: 272 gr3, 4, 5, 6	TCT-DP	cr irr	achiev	fig	fig	oe-i	long gen: gr3-6+
18	He and Wong (2015) Hong Kong	coh long	coh: 3 long: 3 (4,9)	n: 514 nr: 168 gr6	TCT-DP	cr irr	achiev	fig	fig	oe-i	gen: $gr6_{MA}$-$gr6_{MB}$=
19	Hong and Milgram (2010) Israel	coh	2	n study1: 108 gr4, 5, 6 n study2: 71 kg0, 1, 2	TACT	cr	achiev	verb fig	verb fig	oe-f	Study1: flu: gr4-6= Study2: flu: kg1-2=
					ARLPS	cr	achiev	verb	verb	oe-f	Study1: flu: gr4-6+ Study2: flu: kg1-2+
20	Hu, Shi, Han, Wang, and Adey (2010) China	coh	4	n: 1367 nr: 506 gr3, gr4, gr5, gr6	TTCT (part)	cr irr	achiev	fig	verb	oe-f	gen: gr4-gr6+ ori: gr4-gr6+ flu: gr4-gr6+ fle: gr4-gr6+
21	Hui and Lau (2006) Hong Kong	interv-coh	2 (4)	n: 126 nr: 69 gr1, 4	WKCT	cr	achiev	verb fig	verb	oe-f	flu: gr1-gr4−
					TCT-DP	cr	achiev	fig	fig	oe-i	gen: gr1-gr4+
					STT	cr	achiev	fig	verb	oe-i	gen: gr1-gr4+

(*continued on next page*)

(*continued*)

Nr	Author(s) (year) Country	De-sign	Measurements (time between*)	Participants	Abbreviation instrument**	Psychometric characteristics	Type of instrument	Domain stimulus	Domain response	Task taxonomy	Developmental results
22	Jaarsveld et al. (2010) Netherlands	coh	8	n: 511 kg1, 2, gr1, 2, 3, 4, 5, 6	CRT	irr	achiev	fig	fig	oe-i	gen: kg1-(gr)6+
23	Jaarsveld et al. (2012) Germany	coh	4	n: 205 gr1, 2, 3, 4	TCT-DP	irr	achiev	fig	fig	oe-i	gen: gr1-4+
					CRT	irr irr	achiev	fig	fig	oe-i	gen: gr1-4=
24	Jamieson-Proctor and Burnett (2002) Australia	interv-coh	2 (10)	n: 438 nr: 168 gr5/6	CC	cr val	quest	verb	choo	na	gen: gr5/6=
25	Jankowska and Karwowski (2018) Poland	interv-coh	4 (5,10,15)	n: 75 gr1/2, 2/3	TCT-DP	cr	achiev	fig	fig	oe-i	gen: gr1/2-gr2/3+
26	Jastrzebska and Limont (2017) Poland	coh	5	n: 1522 nr: 613 gr2, 3, 4, 5, 6	TCT-DP	cr	achiev	fig	fig	oe-i	gen: gr2-gr6+
27	Kiehn (2003) United States	coh	3	n: 89 gr2, 4, 6	TMC	irr val	achiev	mus	mus	oe-f	gen: gr2-gr6+
28	Kim (2011) United States	coh	7	n: 272.599 nr: kg2, gr1, 2, 3, 4, 5, 6	TTCT (part)		achiev	fig	fig	oe-f	ori: kg-gr6+ flu: kg-gr6+
29	Krampen (2012) Germany, Luxembourg	coh long	study1/2 coh: 5 long: 4 (12,24 36)	n study1: 244 n study2: 312 kg1, 2, gr1, 2, 3, 4, 5, 6	DT	Study1: cr irr Study2: cr irr	achiev	verb fig	verb fig mot	oe-f	Study1: flu: kg1/(gr)3-gr2/6+ fle: kg1/(gr)3-gr2/6+ Study2: flu: kg1/(gr)3-gr2/6+ fle: kg1/(gr)3-gr2/6+
30	Lau and Cheung (2010a, 2010b) Hong Kong	coh	3	n: 2476 nr: 1256 gr4, 5, 6	WKCT (part)	cr	achiev	verb	verb	oe-f	ori: gr4-6= flu: gr4-6+ fle: gr4-6+
					WKCT (part)	cr	achiev	fig	verb	oe-f	ori: gr4-6+ flu: gr4-6+ fle: gr4-6+
31	Leikin (2013) Israel	interv-coh	2 (12)	n: 37 nr: 14 kg0, 1	PMS		achiev	fig	verb	oe-f	gen: kg0-kg1= ori: kg0-kg1= flu: kg0-kg1+ fle: kg0-kg1+
32	Lin and Shih (2016) Taiwan	coh long	coh: 3 long: 2 (12)	n: 92 gr3, 4, 5	CVCTT	irr	achiev	verb	verb	oe-f	gen: gr3-gr5−ori: gr3-gr5− flu: gr3-gr5− fle: gr3-gr5−
					CWRAT	cr	achiev	verb	verb	ce-i	gen: gr3-gr5+
					IPS	cr	achiev	verb	verb	ce-i	gen: gr3-gr5+
33	Maker et al. (2008) United States	coh long	coh: 7 long: 2 (36)	n: 1955 kg2, gr1, 2, 3, 4, 5, 6	TCT-DP		achiev	fig	fig	oe-i	coh gen: kg2-(gr)6+
34	Memmert (2007) Germany	interv-coh	2 (6)	n: 48 nr: 23 gr1	GTS	cr irr	achiev	mot	mot	oe-f	gen: $gr1_{MA}$-$gr1_{MB}$=
35	Memmert (2011) Germany	coh	2	n: 120 nr: gr2, 5	DTT	irr	achiev	fig	fig	oe-f	gen: gr2-gr5+ ori: gr2-gr5= flu: gr2-gr5+ fle: gr2-gr5+
					CT	irr	achiev	mot	verb	oe-f	gen: gr2-gr5+ ori: gr2-gr5+ flu: gr2-gr5+ fle: gr2-gr5+
36	Memmert and Roth (2007) Germany	interv-coh	3 (6,15)	n: 135 nr: 20 gr1/3	GTS	irr	achiev	mot	mot	oe-f	gen: gr1/2-gr2/3=

(*continued on next page*)

(*continued*)

Nr	Author(s) (year) Country	De-sign	Measurements (time between*)	Participants	Abbreviation instrument**	Psychometric characteristics	Type of instrument	Domain stimulus	Domain response	Task taxonomy	Developmental results
37	Mohanty (2015) India	coh	2	n: 80 nr: 40 gr4, 6	CTT		achiev	fig	fig verb	oe-i	gen: gr4-gr6=
38	Mouchiroud and Lubart (2002) France	coh	3	n: 88 gr1, 3, 5	SCT (part)		achiev	verb	verb	oe-f	flu: gr1-5+
					TTCT (part)		achiev	fig	verb	oe-f	ori: gr1-5= flu: gr1-5+
					STT	cr	achiev	fig	verb	oe-i	gen: gr1-5+
39	Mourgues et al. (2016) Saudi Arabia	coh	2	n: 4368 nr: 2740 gr3/5, 6/8	CT	cr	achiev	verb fig	verb fig	oe-f	gen: gr3/5-gr6/8+ ori: gr3/5-gr6/8+ flu: gr3/5-gr6/8+ fle: gr3/5-gr6/8+
40	Ren, Li, Zhang, and Wang (2012) China	coh	3	n: 4162 nr: gr4, 5, 6	TCI	cr	achiev	verb fig	verb fig	oe-f	gen: gr4-6+ ori: gr4-6+ fle: gr4-6+
41	Rostan (2010) United States	coh	2	n: 51 gr4/5, 6/10	DT-LI	cr	achiev	verb fig	fig	oe-i	Results contradictory
42	Russ, Robins, and Christiano (1999) United States	interv-coh	2 (48)	n: 31 gr1/2, 5/6	WKCT (part)		achiev	verb	verb	oe-f	flu: gr1/2-gr5/6= fle: gr1/2-gr5/6=
					APS/AFT	irr	achiev	fig	verb	oe-i	gen: gr1/2-gr5/6=
43	Sak and Maker (2006) United States	coh	5	n: 297 gr1, 2, 3, 4, 5	DA		achiev	math	math	oe-f	gen***: gr1-5+ flu: gr1-5+
44	Sali (2015) Turkey	interv-coh	2 (12)	n: 160 kg2, gr1	TTCT (part)	cr	achiev	fig	fig	oe-f	gen: kg2-gr1= ori: kg2-gr1= flu: kg2-gr1=
45	Sayed and Mohamed (2013) Egypt	coh	8	n: 901 kg1, 2, gr1, 2, 3, 4, 5, 6	TCT-DP	cr	achiev	fig	fig	oe-i	gen: kg1-(gr)6+
46	Subbotsky, Hysted, and Jones (2010) United Kingdom	interv-coh	study1: 2 study2: 2	n study1: 52 kg1, gr1 n study2: 64 gr1, 3	TCAM		achiev	verb fig	mot	oe-f	Study1: ori: kg1-gr1+ flu: kg1-gr1+ Study2: ori: gr1-3= flu: gr1-3=
47	Welter, Jaarsveld, van Leeuwen, and Lachmann (2016) Germany	coh long	long: 2 (10) coh: 4	n: 166 nr: 49 gr1/2/3/4	TCT-DP		achiev	fig	fig	oe-i	Results not applicable
48	Welter, Jaarsveld, and Lachmann (2017) Germany	coh	4	n: 166 gr1, 2, 3, 4	TCT-DP		achiev	fig	fig	oe-i	gen: gr1-4+
					CRT		achiev	fig	fig	oe-i	gen: gr1-4+
49	Wojciehowski and Ernst (2018) United States	interv-coh	2 (8)	n: 75 nr: 11 kg0/gr1	TCAM		achiev	verb fig	mot	oe-f	ori: kg0/gr1$_{MA}$-kg0/gr1$_{MB}$= flu: kg0/gr1$_{MA}$-kg/gr1$_{MB}$=
50	Wu (2010) China	coh	3	n: 357 nr: 180 gr4, 5, 6	CAP (part)	irr	achiev	fig	fig	oe-f	ori: gr4-gr6= flu: gr4-gr6-fle: gr4-gr6=
51	Yeh and Li (2008) Taiwan	coh	3	n: 116 kg1, 2, gr1	PCT	cr irr	achiev	fig	verb fig	oe-f + i	gen: kg1-(gr1)+ ori: kg1-(gr1)+
52	Yi, Hu, Plucker, and McWilliams (2013) China	coh	2	n: 331 nr: 145 gr5, 6	BTCT (part)	cr irr	achiev	verb	verb	oe-f	ori: gr5-gr6= flu: gr5-gr6= fle: gr5-gr6=
					BTCT (part)	cr irr	achiev	fig	fig	oe-f	ori: gr5-gr6= flu: gr5-gr6= fle: gr5-gr6=

Notes: For explanations of used codes, see Table 2 and the text in paragraph 4.2 of the article. *The time between measurements is in months.
**Explanation of the abbreviations of the instruments: AFT = Affect in Fantasy Task, APS = Affect in Play Scale, ARLPS = Ariel Real-Life Problem Solving, ATB = Aurora Test Battery, BTCT=Beijing Test of Creative Thinking, BTMT=Bertsch' Tests of Motor Creativity, CAP=Creativity Assessment Packet, CC=Creativity Checklist, CCRS=The Creative Characteristics Rating Scale, CDAT=Clark's Drawing Abilities Test, CRT=Creative Reasoning Task, CT=The Creativity Test (^{A} and ^{B} are different instruments), CTT=Creative Thinking Test, CVCTT=Chinese Version of Creative Thinking Test, CWRAT=Chinese Word Remote Associates Test, DA = Discover Assessment, DT-LI = Drawing Task for Life or Imagination, DT = Divergent Thinking, DTT = Divergent Thinking Test for Preschool and School children, ET = Evaluative Thinking, FDCT=Franck Drawing Completion Test, GTS = Game Test Situations, IPS=Insight Problem Solving, ITT=Integrative Thinking Task, KRDC=Kriteriji Razvoja Development Criteria, NTCT=New Tests of Creative Thinking, PCT=Preschoolers' Creativity Test, PMS=Pictorial Multiple Solution, SCT=Social Creativity Tasks, STT=Story Telling Task, TACT = Tel Aviv Creativity Test, TCAM = Thinking Creatively in Action and Movement, TCI = Test de Creatividad Infantil, TCT-DP = Test for Creative Thinking- Drawing Production, TMC = Vaughan Test of Musical Creativity, TTCT = Torrance Test of Creative Thinking, WKCT=Wallach Kochan Creativity Tests. (part) = a part of the instrument is used.
***Article 43 (Sak & Maker, 2006) general (creativity) = OFE; originality, flexibility and elaboration together.

Appendix C

Three-level funnel plots of standardized mean differences concerning general creativity, originality, fluency and flexibility

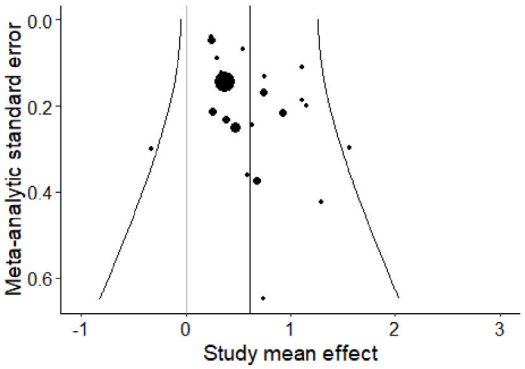


Fig. C1. General creativity

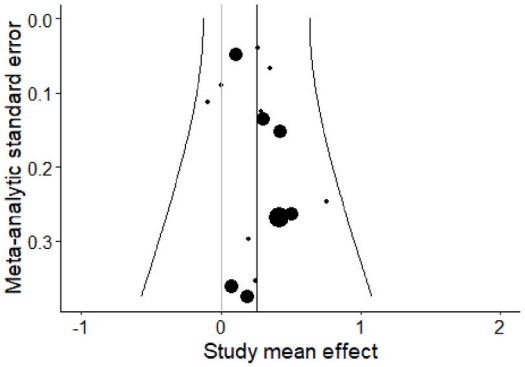


Fig. C2. Originality

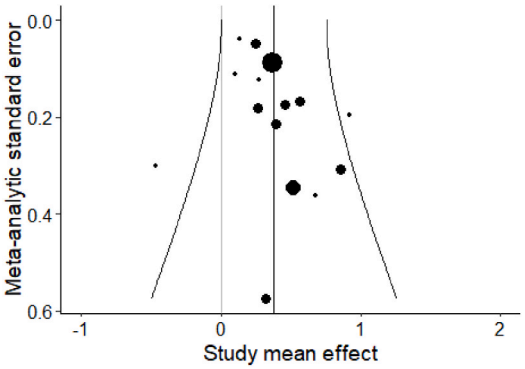


Fig. C3. Fluency

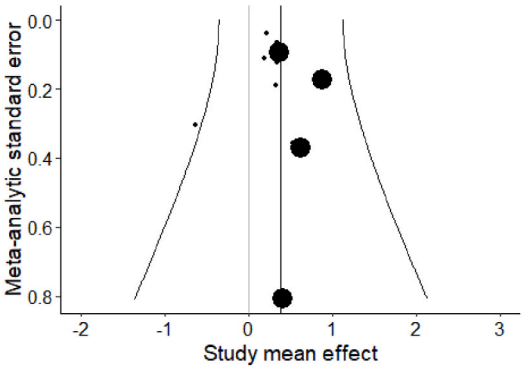


Fig. C4. Flexibility

Note. Size of the dots reflects the number of effect sizes in a study. See Fernández-Castilla et al. (2020) for detailed explanations.

Appendix D

Three-level forest plot of standardized mean differences

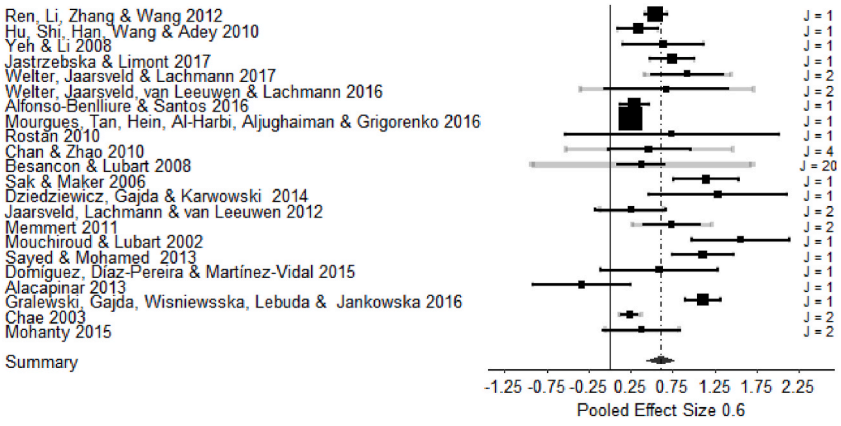


Fig. D1. General creativity

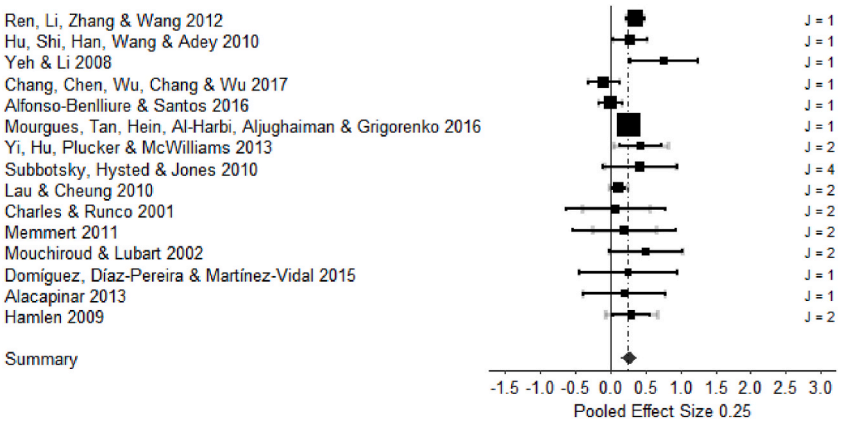


Fig. D2. Originality

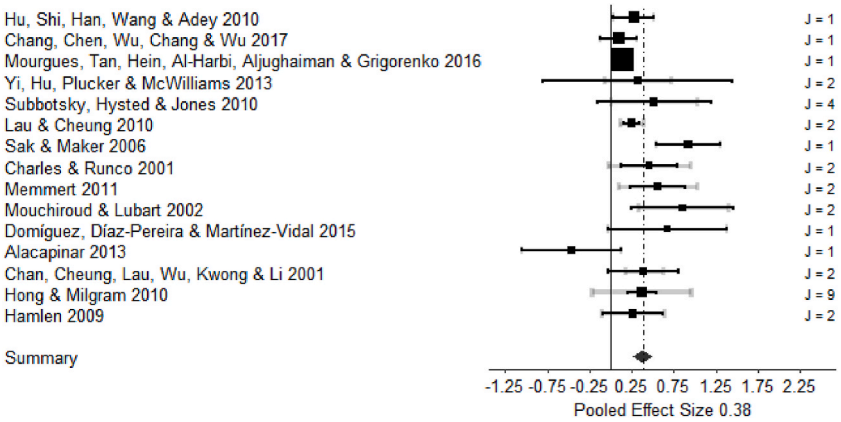


Fig. D3. Fluency

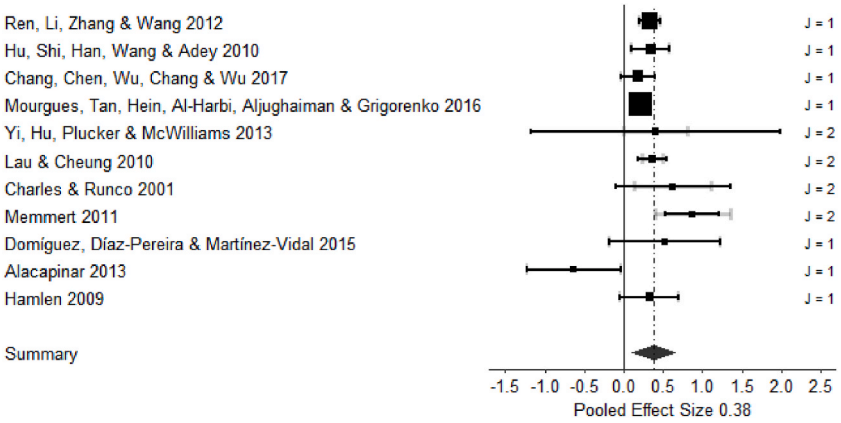


Fig. D4. Flexibility

Note: J indicates the number of effect sizes in the study. Black squares represent the (average) effect size in a study. Size of the square reflects the weight in the meta-analysis. Black confidence intervals represent 95% confidence intervals around the study's average effect size. Grey confidence intervals are based on the median standard error associated with the study's multiple effect sizes. See Fernández-Castilla et al. (2020) for detailed explanations.

References

Abdulla, A. M., Paek, S. H., Cramond, B., & Runco, M. A. (2018). Problem finding and creativity: A meta-analytic review. *Psychology of Aesthetics, Creativity, and the Arts, 16*(1), 3–14. https://doi.org/10.1037/aca0000194

Alacapinar, F. G. (2013). Grade level and creativity. *Egitim Arastirmalari- Eurasian journal of Educational Research, 50*, 247–266. https://eric.ed.gov/?id=EJ1059915.

Alfonso-Benlliure, V., & Santos, M. R. (2016). Creativity development trajectories in elementary education: Differences in divergent and evaluative skills. *Thinking Skills and Creativity, 19*, 160–174. https://doi.org/10.1016/j.tsc.2015.11.003

Baer, J. (1998). The case for domain specificity in creativity. *Creativity Research Journal, 11*(2), 173–177. https://doi.org/10.1207/s15326934crj1102_7

Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roeper Review, 27*(3), 158–163. https://doi.org/10.1080/02783190509554310

Baer, J., & Kaufman, J. C. (2017). The amusement park theoretical model of creativity. In J. C. Kaufman, V. P. Glăveanu, & J. Baer (Eds.), *The Cambridge handbook of Creativity across domains* (pp. 8–17). Cambridge University Press.

Barbot, B. (2019). Measuring creativity change and development. *Psychology of Aesthetics, Creativity, and the Arts, 13*(2), 203–210. https://doi.org/10.1037/aca0000232

Barbot, B., Besançon, M., & Lubart, T. I. (2011). Assessing creativity in classroom. *The Open Education Journal, 4*(1), 58–66. https://hal.parisnanterre.fr/hal-01392552.

Barbot, B., Besançon, M., & Lubart, T. I. (2016). The generality-specificity of creativity: Exploring the structure of creative potential with EPoC. *Learning and Individual Differences, 52*, 178–187. https://doi.org/10.1016/j.lindif.2016.06.005

Beghetto, R. A. (2016). Creative openings in the social interactions of teaching. *Creativity: Theories-Research-Applications, 3*, 261–273. https://doi.org/10.1515/ctra-2016-0017

Besançon, M., & Lubart, T. (2008). Differences in the development of creative competencies in children schooled in diverse learning environments. *Learning and Individual Differences, 18*(4), 381–389. https://doi.org/10.1016/j.lindif.2007.11.009

Besançon, M., Lubart, T. I., & Barbot, B. (2013). Creative giftedness and educational opportunities. *Giftedness in education*. https://hal.parisnanterre.fr//hal-01392541.

Chae, S. (2003). Adaption of a picture-type creativity test for pre-school children. *Language Testing, 20*(2), 178–188. https://doi.org/10.1191/0265532203lt251oa

Chan, D. W., Cheung, P.-C., Lau, S., Wu, W. Y. H., Kwong, J. M. L., & Li, W.-L. (2001). Assessing ideational fluency in primary students in Hong Kong. *Creativity Research Journal, 13*(3–4), 359–365. https://doi.org/10.1207/S15326934crj1334_13

Chang, Y.-L., Chen, H.-C., Wu, I.-C., Chang, J.-H., & Wu, C.-L. (2017). Developmental trends of divergent thinking and feeling across different grades for Taiwanese adolescence between 1990's and 2010's. *Thinking Skills and Creativity, 23*, 112–128. https://doi.org/10.1016/j.tsc.2016.12.002

Chan, D. W., & Zhao, Y. (2010). The relationship between drawing skill and artistic creativity: Do age and artistic involvement make a difference? *Creativity Research Journal, 22*(1), 27–36. https://doi.org/10.1080/10400410903579528

Charles, R. E., & Runco, M. A. (2001). Developmental trends in the evaluative and divergent thinking of children. *Creativity Research Journal, 13*(3), 417–437. https://doi.org/10.1207/S15326934crj1334_19

Cheung, P. C., Lau, S., Chan, D. W., & Wu, W. Y. H. (2004). Creative potential of school children in Hong Kong: Norms of the wallach-kogan creativity tests and their implications. *Creativity Research Journal, 16*(1), 69–78. https://doi.org/10.1207/s15326934crj1601_7

Claxton, A. F., Pannells, T. C., & Rhoads, P. A. (2005). Developmental trends in the creativity of school-age children. *Creativity Research Journal, 17*(4), 327–335. https://doi.org/10.1207/s15326934crj1704_4

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal, 18*, 391–404. https://doi.org/10.1207/s15326934crj1803_13

Darvishi, Z., & Pakdaman, S. (2012). Fourth grade slump in creativity: Development of creativity in primary school children. *GSTF International Journal of Law and Social Sciences, 1*, 40–48. https://doi.org/10.1037/e527322013-009

Davies, D., Jindal-Snape, D., Collier, Digby, C., Hay, P., & Howe, A. (2013). Creative learning environments in education - a systematic literature review. *Thinking Skills and Creativity, 8*, 80–91. https://doi.org/10.1016/j.tsc.2012.07.004

Domíguez, A., Díaz-Pereira, M. P., & Martínez-Vidal, A. (2015). The evolution of motor creativity during primary education. *Journal of Human Sport and Exercise, 10*(2), 583–591. https://doi.org/10.14198/jhse.2015.102.05

Duh, M., & Budefeld, A. (2018). Gender difference in the linear relationship among factors of drawing-related creativity in second-grade primary school students. *Nova Prisutnost, 16*(2), 351–367. https://doi.org/10.31192/np.16.2.9

Dziedziewicz, D., Gajda, A., & Karwowski, M. (2014). Developing children's intercultural competence and creativity. *Thinking Skills and Creativity, 13*, 32–42. https://doi.org/10.1016/j.tsc.2014.02.006

Dziedziewicz, D., Oledzka, D., & Karwowski, M. (2013). Developing 4-to 6-year-old children's figural creativity using a doodle-book program. *Thinking Skills and Creativity, 9*, 85–95. https://doi.org/10.1016/j.tsc.2012.09.004

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods, 52*(5), 2031–2052. https://doi.org/10.3758/s13428-020-01373-9

Francom, G. F. (2017). Principles for task-centered instructions. In C. M. Reigelth, B. J. Beatty, & R. D. Myers (Eds.), *Instructional-Design theories and models, IV, the learner-centered paradigm of education*. Routledge Yaylor & Francis Group.

Fullan, M., & Langworthy, M. (2014). *A rich seam: How new pedagogies find deep learning*. Pearson.

Gajda, A., Karwowski, M., & Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology, 109*(2), 269–299. https://doi.org/10.1037/edu0000133

Glăveanu, V. P. (2013). Rewriting the language of creativity: The five A's framework. *Review of General Psychology, 17*(1), 69–81. https://doi.org/10.1037/a0029528

Glăveanu, V. P. (2014). *Distributed creativity: Thinking outside the box of the creative individual*. Springer.

Glăveanu, V. P., & Kaufman, K. C. (2019). Creativity: A historical perspective. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The cambridge handbook of creativity* (2nd ed.). Cambridge University Press.

Gralewski, J., Gajda, A., Wisniewsska, E., Lebuda, I., & Jankowska, D. M. (2016). Slumps and jumps: Another look at developmental changes in creative abilities. *Creativity: Theories-Research-Applications, 3*(1), 152–177. https://doi.org/10.1515/ctra-2016-0011

Gubbels, J., Segers, E., & Verhoeven, L. (2017). Predicting the development of analytical and creative abilities in upper elementary grades. *Creativity Research Journal, 29*(4), 433–441. https://doi.org/10.1080/10400419.2017.1376548

Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444–454.

Guilford, J. P. (1951). *A factor-analytic study of creative thinking. Hypothesis and description of tests* (p. 1). Reports from the Psychological Laboratory, the University of Southern California.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Hamlen, K. R. (2009). Relationships between computer and video game play and creativity among upper elementary school students. *Journal of Educational Computing Research, 40*(1), 1–21. https://doi.org/10.2190/EC.40.1.a

He, W.-J. (2018). A 4-year longitudinal study of the sex-creativity relationship in childhood, adolescence, and emerging adulthood: Findings of mean and variability analyses. *Frontiers in Psychology, 9*, 2331. https://doi.org/10.3389/fpsyg.2018.02331

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426–445. https://doi.org/10.1037/1082-989X.9.4.426

He, W.-J., & Wong, W.-C. (2015). Creativity slump and school transition stress: A sequential study from the perspective of the cognitive-relational theory of stress. *Learning and Individual Differences, 43*, 185–190. https://doi.org/10.1016/j.lindif.2015.08.034

Hong, E., & Milgram, R. M. (2010). Creative thinking ability: Domain generality and specificity. *Creativity Research Journal, 22*(3), 272–287. https://doi.org/10.1080/10400419.2010.503535

Huang, P. S., Chen, H. C., & Liu, C. H. (2012). The development of Chinese word remote associates test for college students. *Psychological Testing, 59*, 581–607.

Hui, A., & Lau, S. (2006). Drama education: A touch of the creative mind and communicative-expressive ability of elementary school children in Hong Kong. *Thinking Skills and Creativity, 1*, 34–40. j.tsc.2005.06.001.
Hu, W., Shi, Q. Z., Han, Q., Wang, X., & Adey, P. (2010). Creative scientific problem finding and its developmental trend. *Creativity Research Journal, 22*(1), 46–52. https://doi.org/10.1080/10400410903579551
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*(3), 304–319. https://doi.org/10.1080/10400419.2010.503541
Jaarsveld, S., Lachmann, T., & van Leeuwen, C. (2012). Creative reasoning across developmental levels: Convergence and divergence in problem creation. *Intelligence, 40*(2), 172–188. https://doi.org/10.1016/j.intell.2012.01.002
Jamieson-Proctor, R., & Burnett, P. C. (2002). Elementary students, creativity, and technology: Investigation of an intervention designed to enhance personal creativity. *Computers in the Schools, 19*(1), 33–48. https://doi.org/10.1300/J025v19n01_04
Jankowska, D. M., & Karwowski, M. (2018). Family factors and development of creative thinking. *Personality and Individual Differences, 142*, 202–206. https://doi.org/10.1016/j.paid.2018.07.030
Jastrzebska, D., & Limont, W. (2017). Not only jumps, slumps, but also mini plateau. creative potential assessed by the test for creative thinking-drawing production. A cross-sectional study of polish students aged from 7 to 18. *Creativity Research Journal, 29*, 337–342. https://doi.org/10.1080/10400419.2017.1360060
Kaufman, J. C., & Baer, J. (2005). The amusement park theory of creativity. In J. C. Kaufman, & J. Baer (Eds.), *The Cambridge handbook of Creativity across domains* (pp. 321–328). Erlbaum.
Kiehn, M. T. (2003). Development of music creativity among elementary school students. *Journal of Research in Music Education, 51*(4), 278–288. https://doi.org/10.2307/3345655
Kim, K. H. (2011). The creativity crisis: The decrease in creative thinking scores on the Torrance tests of creative thinking. *Creativity Research Journal, 23*(4), 285–295. https://doi.org/10.1080/10400419.2011.627805
Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods, 2*(1), 61–76. https://doi.org/10.1002/jrsm.35
Krampen, G. (2012). Cross-sequential results on creativity development in childhood within two different school systems: Divergent performances in Luxemburg versus German kindergarten and elementary school students. *Europe's Journal of Psychology, 8*(3), 423–448. https://doi.org/10.5964/ejop.v8i3.468
Kupers, E., Lehmann-Wermser, A., McPherson, G., & Van Geert, P. (2019). Children's creativity: A theoretical framework and systematic review. *Review of Educational Research, 89*, 93–124. https://doi.org/10.3102/0034654318815707
Lau, S., & Cheung, P. C. (2010a). Creativity assessment: Comparability of the electronic and paper-and-pencil versions of the wallach–kogan creativity tests. *Thinking Skills and Creativity, 5*, 101–107. https://doi.org/10.1016/j.tsc.2010.09.004
Lau, S., & Cheung, P. C. (2010b). Developmental trends of creativity: What twists of turn do boys and girls take at different grades? *Creativity Research Journal, 22*(3), 329–336. https://doi.org/10.1080/10400419.2010.503543
Leikin, M. (2013). The effect of bilingualism on creativity: Developmental and educational perspectives. *International Journal of Bilingualism, 17*(4), 431–447. https://doi.org/10.1177/1367006912438300
Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *BMJ, 339*, b2700. https://doi.org/10.1136/bmj.b2700, 2009.
Lin, W.-L., & Shih, Y.-L. (2016). The developmental trends of different creative potentials in relation to children's reasoning abilities: From a cognitive theoretical perspective. *Thinking Skills and Creativity, 22*, 36–47. https://doi.org/10.1016/j.tsc.2016.08.004
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications Inc.
Lowenfeld, V. (1960). *Von wesen schöpferischen Gestaltens*. Europäische Verlagsanstalt.
Lubart, T. I., Besançon, M., & Barbot, B. (2011). *Évaluation du potentiel créatif des enfants*. Paris: EPoC.
Lubart, T. I., & Georgsdottir, A. (2004). Creativity: Developmental and cross-cultural issues. In S. Lau, A. H. H. Hui, & G. Y. C. Ng (Eds.), *Creativity: When east meets west* (pp. 23–54). World Scientific Publishing Company. https://doi.org/10.1142/9789812567192_0003.
Lubart, T. I., & Lautrey, J. (1995). *Relationships between creative development and cognitive development [Paper presentation]*. Krakow, Poland: *Europe*an Conference on Developmental Psychology, 7th Biennial Meeting.
Maker, C. J., Jo, S., & Muammar, O. M. (2008). Development of creativity: The influence of varying levels of implementation of the discover curriculum model, a non-traditional pedagogical approach. *Learning and Individual Differences, 18*(4), 402–417. https://doi.org/10.1016/j.lindif.2008.03.003
Memmert, D. (2007). Can creativity be improved by an attention-broadening training program? *Creativity Research Journal, 19*(2–3), 281–291. https://doi.org/10.1080/10400410701397420
Memmert, D. (2011). Creativity, expertise, and attention: Exploring their development and their relationships. *Journal of Sports Sciences, 29*(1), 93–102. https://doi.org/10.1080/02640414.2010.528014
Memmert, D., & Roth, K. (2007). The effects of non-specific and specific concepts on tactical creativity in team ball sports. *Journal of Sports Sciences, 25*(12), 1423–1432. https://doi.org/10.1080/02640410601129755
Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research & Development, 50*(3), 43–59. https://doi.org/10.1007/BF02505024
Merrill, M. D. (2007). A task-centered instructional strategy. *Journal of Research on Technology in Education, 40*(1), 5–22. https://doi.org/10.1080/15391523.2007.10782493
Mohanty, A. (2015). Information processing and creative thinking abilities of residential and non-residential school children: A pilot study. *Sage Open, 5*(4), 1–12. https://doi.org/10.1177/2158244015611452
Mouchiroud, C., & Lubart, T. (2002). Social creativity: A cross-sectional study of 6- to 11-year-old children. *International Journal of Behavioral Development, 26*(1), 60–69. https://doi.org/10.1177/016502540202600111
Mourgues, C. V., Tan, M., Hein, S., Al-Harbi, K., Aljughaiman, A., & Grigorenko, E. L. (2016). The relationship between analytical and creative cognitive skills from middle childhood to adolescence: Testing the threshold theory in the kingdom of Saudi Arabia. *Learning and Individual Differences, 52*, 137–147. https://doi.org/10.1016/j.lindif.2015.05.005
Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology, 21*(1), 34–77. https://doi.org/10.1080/10463281003765323
Piaget, J. (1951). Principle factors determining intellectual evolution from childhood to adult life. In D. Rapaport (Ed.), *Organization and pathology of thought*. Columbia University Press. https://doi.org/10.7312/rapa92214-009 (Original work published in 1937).
Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*, 83–96. https://doi.org/10.1207/s15326985ep3902_1
Preiss, D. D., Grau, V., Ortiz, D., & Bernardino, M. (2016). What do we know about the development of creativity in south America? In D. D. Preiss (Ed.), *New directions of child and adolescent development: 152. Child and adolescent development in Latin America* (pp. 85–97). https://doi.org/10.1002/cad.20157
R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org.
Raven, J., Raven, J. C., & Court, J. H. (1998a). *Manual for Raven's progressive matrices and vocabulary scales. Section 1: General overview*. Oxford, UK: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.
Raven, J., Raven, J. C., & Court, J. H. (1998b). *Manual for Raven's progressive matrices and vocabulary scales. Section 2: The coloured progressive matrices*. Oxford, UK: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.
Raven, J., Raven, J. C., & Court, J. H. (1998c). *Manual for Raven's progressive matrices and vocabulary scales. Section 4: The advanced progressive matrices*. Oxford, UK: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.
Raven, J., Raven, J. C., & Court, J. H. (1998d). *Manual for Raven's progressive matrices and vocabulary scales. Section 5: The Mill Hill vocabulary scale*. Oxford, UK: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.

Ren, F., Li, X., Zhang, H., & Wang, L. (2012). Progression of Chinese students' creative imagination from elementary through high school. *International Journal of Science Education, 34*(13), 2043–2059. https://doi.org/10.1080/09500693.2012.709334
Rosenblatt, E., & Winner, E. (1988). The art of children's drawings. *Journal of Aesthetic Education, 22*, 3–15. https://doi.org/10.2307/3332960
Rostan, S. M. (2010). Studio learning: Motivation, competence, and the development of young art students' talent and creativity. *Creativity Research Journal, 22*(3), 261–271. https://doi.org/10.1080/10400419.2010.503533
Runco, M. A. (2008). Divergent thinking is not synonymous with creativity. *Psychology of Aesthetics, Creativity, and the Arts, 2*(2), 93–96. https://doi.org/10.1037/1931-3896.2.2.93
Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal, 24*(1), 66–75. https://doi.org/10.1080/10400419.2012.652929
Runco, M. A., & Acar, S. (2019). Divergent thinking. In r. J. Sternberg (Ed.), *Handbook of creativity* (2nd ed., pp. 224–251). Cambridge University Press.
Russ, S. W., Robins, A. L., & Christiano, B. A. (1999). Pretend play: Longitudinal prediction of creativity and affect in fantasy in children. *Creativity Research Journal, 12*(2), 129–139. https://doi.org/10.1207/s15326934crj1202_5
Said-Metwaly, S., Fernández-Castilla, B., Kyndt, E., Van den Noortgate, W., & Barbot, B. (2020). Does the fourth-grade slump in creativity actually exist? A meta-analysis of the development of divergent thinking in school-age children and adolescents. *Educational Psychology Review, 33*, 275–298. https://doi.org/10.1007/s10648-020-09547-9
Sak, U., & Maker, C. J. (2006). Developmental variation in children's creative mathematical thinking as a function of schooling, age, and knowledge. *Creativity Research Journal, 18*(3), 279–291. https://doi.org/10.1207/s15326934crj1803_5
Sali, G. (2015). A longitudinal study on the development of creativity in children. *The Anthropologist, 20*(1,2), 93–100. https://doi.org/10.1080/09720073.2015.11891728
Sayed, E. M., & Mohamed, A. H. H. (2013). Gender differences in divergent thinking: Use of the test of creative thinking-drawing production on an Egyptian sample. *Creativity Research Journal, 25*(2), 222–227. https://doi.org/10.1080/10400419.2013.783760
Saywer, R. K. (2012). *Explaining creativity: The science of human innovation* (2nd ed.). Oxford University Press.
Scott, G., Leritz, L. E., & Mumford, M. D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal, 16*, 361–388. https://doi.org/10.1080/10400410409534549
Simonton, D. K., & Damian, R. I. (2013). Creativity. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology* (pp. 795–807). Oxford University Press.
Smith, G. F. (1998). Idea-generation techniques: A formulary of active ingredients. *Journal of Creative Behavior, 32*(2), 107–133. https://doi.org/10.1002/j.2162-6057.1998.tb00810.x
Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 3–15). Cambridge University Press.
Subbotsky, E., Hysted, C., & Jones, N. (2010). Watching films with magical content facilitates creativity in children. *Perceptual and Motor Skills, 111*(1), 261–277. https://doi.org/10.2466/04.09.11.PMS.111.4.261-277
Torrance, E. P. (1966). *The torrance tests of creative thinking: Norms-technical manual.* Personal Press.
Torrance, E. P. (1967). *Understanding the fourth grade slump in creative thinking (Report No. BR-5-0508; CRP994)*. U. S. Office of Education.
Torrance, E. P. (1968). A longitudinal examination of the fourth-grade slump in creativity. *Gifted Child Quarterly, 12*, 195–199. https://doi.org/10.1177/001698626801200401
Torrance, E. P. (1981). *Thinking creatively in action and movement*. Scholastic Testing Service.
Torrance, E. P. (1990). *The Torrance tests of creative thinking norms - technical manual figural (streamlined) forms A & B*. Bensenville, IL: Scholastic Testing Service, Inc.
Tsamir, P., Tirosh, D., Tabach, M., & Leverson, E. (2009). Multiple solution methods and multiple outcomes—is it a task for kindergarten children? *Educational Studies in Mathematics, 73*, 217–231. https://doi.org/10.1007/s10649-009-9215-z
Urban, K. K. (2004). Assessing creativity: The test for creative thinking - drawing production (TCT-DP): The concept, application, evaluation, and international studies. *Psychology Science, 46*(3), 387–397. https://psycnet.apa.org/record/2005-05211-008.
Urban, K. K., & Jellen, H. G. (1996). *Test for creative thinking - drawing production (TCTDP). Swets and zeitlinger*.
Van Merrienboer, J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Taylor & Francis inc.
Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*(2), 576–594. https://doi.org/10.3758/s13428-012-0261-6
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03
Wakefield, J. F. (1987). *The outlook for creativity tests*. Orlando, FL, United States: Council for Exceptional Children Topical Conference on the Future of Special Education [Paper presentation] https://eric.ed.gov/?id=ED292249 [Paper presentation].
Wakefield, J. F. (1989). *March 27-31). An arts orientation, cognitive skills and creative behavior*. San Francisco, CA, United States: American Educational Research Association, 70th Annual Meeting [Paper presentation] https://eric.ed.gov/?id=ED310045.
Wakefield, J. F. (1991). The outlook for creativity tests. *Journal of Creative Behavior, 25*, 184–193. https://eric.ed.gov/?id=ED292249.
Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity intelligence distinction*. Holt, Rinehart, & Winston. https://doi.org/10.1001/archpsyc.1966.01730090105021
Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg, & J. E. Davidson (Eds.), *The nature of insight* (pp. 157–196). Cambridge (MA): MIT Press. https://doi.org/10.7551/mitpress/4879.003.0009.
Welter, M., Jaarsveld, S., & Lachmann, T. (2017). Problem space matters: The development of creativity and intelligence in primary school children. *Creativity Research Journal, 29*(2), 125–132. https://doi.org/10.1080/10400419.2017.1302769
Welter, M. M., Jaarsveld, S., van Leeuwen, C., & Lachmann, T. (2016). Intelligence and creativity: Over the threshold together? *Creativity Research Journal, 28*(2), 212–218. https://doi.org/10.1080/10400419.2016.1162564
Williams, F. E. (1980). *Creativity assessment packet: (CAP)*. D.O.K. Publishers.
Williams, F. E. (1993). *Cap: Creativity assessment packet*. Austin, TX: Pro-Ed.
Wojciehowski, M., & Ernst, J. (2018). Creative by nature: Investigating the impact of nature preschools on young children's creative thinking. *International Journal of Early Childhood Environmental Education, 6*(1), 3–20. https://eric.ed.gov/?id=EJ1193490.
Wu, W. (2010). Developmental trend study of divergent thinking among students from primary to middle school. *International Journal of Psychological Studies, 2*(1), 122–127. https://doi.org/10.5539/ijps.v2n1p122
Wu, C.-L., & Cheng, H.-C. (2019). The influence of creativity on incongruity-resolution and nonsense humor comprehension. *Creativity Research Journal, 31*(1), 110–118. https://doi.org/10.1080/10400419.2019.1577675
Yeh, Y.-C., & Li, M.-L. (2008). Age, emotion regulation strategies, temperament, creative drama, and preschoolers' creativity. *Journal of Creative Behavior, 42*(2), 131–148. https://doi.org/10.1002/j.2162-6057.2008.tb01291.x
Yi, X., Hu, W., Plucker, J. A., & McWilliams, J. (2013). Is there a developmental slump in creativity in China? The relationship between organizational climate and creativity development in Chinese adolescents. *Journal of Creative Behavior, 47*(1), 22–40. https://doi.org/10.1002/jocb.21
Zyga, O., Ivcevic, Z., Hoffmann, J., & Palomera, R. (2021). *Developmental trends in creative ability: A cross-sectional examination of figural and verbal domains across the school-age years. Psychology of aesthetics, creativity, and the arts*. Advance online publication. https://doi.org/10.1037/aca0000425