# Developmental Trends in Creative Ability:

# A Cross-Sectional Examination of Figural and Verbal Domains Across the School-Age Years

Abstract

Studies of creative ability across ages and grade levels show inconsistent results. Further, past studies have rarely used tasks from multiple creativity domains. The current study examined the differences in creative ability across grade levels ($1^{st}$-$10^{th}$ grade), task domains (verbal and figural), and performance criteria (creativity, novelty/originality, elaboration, emotion use). Results showed that (1) within task domain, ratings of creativity tended to steadily increase with grade level, while other performance criteria showed slumps and jumps; (2) more jumps and slumps were evident in the verbal task than in the figural task, and (3) correlations between different performance criteria on the figural and verbal tasks is similar across grade levels and speaks to domain specificity of creative abilities. Taken together, these results point to a complex relationship between grade level, task domain, and creativity-related performance criteria, and the need for more research across these areas.

*Keywords: divergent thinking, verbal creativity domain, figural creativity domain, child development*

# Developmental Trends in Creative Ability: A Cross-Sectional Examination of Figural and Verbal Domains Across the School-Age Years

Creative thinking skills are critical 21st-century skills (Dede, 2010). In a report on the future of jobs, the World Economic Forum identified creativity, ideation, and complex problem solving as key skills for the changing economy (World Economic Forum, 2020). Scholars across disciplines also point out to the importance of creativity for a broad range of outcomes, such as student performance, individual achievement, health and well-being, and communication and interpersonal skills (Forgeard & Kaufman, 2016). Age and school grade differences in creative ability thus become important for the theoretical reasons of understanding developmental trends, as well as for practical reasons, such as informing needs for intervention at different points during schooling.

Torrance (1967, 1968) observed a 4$^{th}$ grade slump in divergent thinking test performance and his observation has spurred decades of research. A recent meta-analysis of performance on divergent thinking tests in children and adolescents however, did not find evidence for the famed 4$^{th}$ grade slump, but rather found evidence for a 7$^{th}$ grade slump instead (Said-Metwaly, Fernández-Castilla, Kyndt, Van den Noortgate, & Barbot, 2020). The meta-analysis pointed to the importance of moderators, such as test used, content domain (e.g., verbal vs. figural), and country. Studies have also evaluated creative ability across tasks using multiple performance criteria, such as originality, flexibility (breadth of categories of ideas), and fluency (number of different ideas produced) in individuals ranging from preschool age and into older adulthood (e.g., Camp, 1994; Gralewski, Gajda, Wisniewska, Lebuda, & Jankowska, 2016; Jastrzębska & Limont, 2017; Krampen, 2012; Runco, 1999). Overall, the consistent takeaway has been that the

development of creative thinking shows slumps, jumps, and bumps across age (Charles & Runco, 2000-2001; Claxton, Pannells, & Rhoads, 2005; Jastrzebska & Limont, 2017; Klaus, 1991; Kleibeuker, Dreu, & Crone, 2013); however, more work needs to be done on understanding these jumps and bumps across age or school grades, countries, and domains of creativity. In this study, we examine performance on verbal and figural creativity tasks across primary and secondary school grades in Spain, and examine four performance criteria – creativity, novelty, elaboration, and emotion use.

**Jumps, Bumps, and Slumps**

Research has shown that there are slumps and jumps in creative ability across age and/or school grade levels (Camp, 1994; Gralewski et al., 2016; Jastrzębska & Limont, 2017; Krampen, 2012; Runco, 1999). Torrance (1967, 1968) identified a 4th grade slump across seven cultures, which inspired subsequent research on developmental trends in creative ability. Support for the 4th grade slump has been inconsistent, however, with some studies supporting Torrance's findings (e.g., Darvishi & Pakdaman, 2012; Lubart & Georgsdottir, 2004) and others not finding support for a dip in performance on creative thinking tests in 4th grade (e.g., Gralewski, Gajda, Wiśniewska, Lebuda, Jankowska, 2016; Jastrzębska, & Limont, 2017; Lin & Shih, 2016; Sak & Maker, 2006).

A recent meta-analysis of 41 studies examining divergent thinking test performance from grades 1-12, found general upward trends for divergent thinking across grade levels, with some discontinuities (Said-Metwaly, Fernández-Castilla, Kyndt, Van den Noortgate, & Barbot, 2020). Specifically, divergent thinking was higher at each successive grade level up until a dip in 7th grade, reached a peak in 9th grade, with further dips in grades 10 and 11. While the meta-analysis did not find evidence for a general 4th grade slump (Said-Metwaly et al., 2020), moderator

analyses showed that whether a 4th grade slump is identified depends on the test and creativity domain, country, and giftedness. The 7th grade slump was detected across performance criteria and it was also moderated by test type, content domain, and gender.

The meta-analysis found sufficient data in previous research to analyze three indicators of creative thinking – fluency, flexibility, and originality – and found that each followed similar upward trajectories, though with some notable exceptions (Said-Metwaly et al., 2020). Fluency scores between grades 9 and 10 held steady rather than declining, with the drop in scores not appearing until grade 11. Flexibility scores on the other hand, significantly increased in grade 10 and plateaued in grade 11 rather than dropping at all. Finally, originality was the only indicator to show a drop in scores in grade 3, and, similar to flexibility, had a rise in grade 10. Moderator analyses showed that each divergent thinking indicator had varied and different interactions with other factors, such as gender, intellectual giftedness, culture, and the divergent thinking test administered. For example, the change in originality scores from third to fourth grade were in the negative direction for the Wallach-Kogan Creativity Tests, and in the positive direction for the Torrance Tests of Creative Thinking. Furthermore, content domain (verbal or figural) was a moderator of 7th grade fluency and flexibility scores, with a larger decline from 6th to 7th grade for the figural domain than the verbal domain.

Few studies have measured age or grade difference in additional creative thinking indicators, such as elaboration or emotion use. Little research exists about developmental trends in elaboration, though research based on six normative samples for the Torrance Test of Creative Thinking - Figural Version suggests that elaboration increases linearly from kindergarten to 5th grade, stabilizes in 6th grade, and then increases throughout high school in the United States (which typically includes 9th through 12th grades; Kim, 2011). A longitudinal study using the

figure completion task also found an increase in elaboration from 4th to 6th and to 9th grade (Claxton et al., 2005). Even fewer studies have examined affective expression in creative thinking tasks. In a large scale study including participants from kindergarten to college, Gralewski et al. (2017) administered the Test of Creative Thinking – Drawing Production, which was scored for use of humor and affectivity. This study found that use of humor and affectivity in drawings increased steadily with age with no significant slumps. Because creativity in the expressive domains (e.g., art, writing) includes expression and communication of emotion (Botella, Glaveanu, Zenasni, Storme, Myszkowski, Wolff, & Lubart, 2013; Glaveanu et al., 2013; Pelowski, Specker, Gerger, Leder, & Weingarden, 2020), it is important to include measures of emotion use and expression in studies of developmental trends in performance on creative thinking tests.

**Factors Affecting Slumps, Bumps, and Jumps Across Development**

Differences in observed performance on creative thinking tests across ages and school grade levels can be related to three (related) groups of factors: (1) environmental factors, especially pertaining to school transitions; (2) asynchronicity in development; and (3) task specificity and methodological issues (Barbot, Lubart, & Besançon, 2016).

Torrance (1963) was the first to propose that environmental factors, including stress-inducing experiences in transitioning through educational levels, could impact creative ability. A need for acceptance by peers and an emphasis placed on meeting compulsory school expectations may cause students to focus more on appropriate or conventional behaviors rather than exploring novel ways of thinking or behaving, especially in early adolescence, when fitting into a peer group can be especially salient (Adler & Adler, 1995; Cropley, 2001; Lau & Cheng, 2010; Runco & Charles, 1997). Further, at transitional time points (e.g., from primary to

secondary school), there is a focus on building convergent thinking skills. Given these demands, certain grade-level differences in creative performance should be evident, such as decreases or slumps in performance during these transitional points (Runco & Charles, 1997; Smith & Carlsson, 1990; Torrance, 1963), just as negative effects of school transitions have been found for academic motivation and achievement, as well as personal and social well-being (e.g., from elementary to middle school and from middle school to high school; Andrew & Flashman, 2017; Eccles et al., 1993; Gajda, Karwowski, & Beghetto, 2017).

Cross-cultural differences in child rearing and education systems are another potential environmental factor impacting grade-related differences (Dahlman et al., 2013; Lubart & Georgsdottir, 2004), though not the main focus of this study, which takes place within a single country. As societies place varying emphasis on compulsory schooling (largely convergent thinking-oriented) versus exploration of the arts, individuality, and creative development (Lubart & Georgsdottir, 2004), the points at which creative slumps or jumps might occur can vary. As noted by Barbot and colleagues (2016), however, the potential impact is complex; when considering context, one must consider its intersection with the individual, citing the somewhat counterintuitive concept of "diversifying experiences" – in which adversity becomes a potential creativity catalyst (e.g., Damian & Simonton, 2015; Gocłowska, Damian, & Mor, 2018). Indeed, this need to better understand the intersection of test, domain, test taker, and country is again underlined by the recent meta-analysis of Said-Metwaly et al. (2020).

To extract the most information from the available studies in the literature, Said-Metwaly et al. (2020) meta-analysis combined countries of study in three broad cultural groups: Western (e.g., United States, Western European countries), Middle East (including Arab countries) and Eastern (including Asian countries). The resulting analyses could not address whether the timing

of slumps or bumps is related to timing of school transitions, which vary across countries, as well as sometimes within the same country (such as in the United States). A cross-sequential study that followed five cohorts of kindergarten and elementary school students for three years in Luxemburg and Germany points to the importance of analyses that account for specific times of transitions (Krampen, 2012). In both countries performance on creative thinking tests increased up to the next to last grade before the transition (up to 5th grade in Luxemburg where transition happens after 7th grade and up to 3rd grade in Germany where transition happens in 5th grade).

Asynchronicity in development is another factor that has been theoretically proposed to influence timing of slumps and bumps in performance on creative thinking tests is asynchronicity in development (Barbot et al., 2016). Because not all aspects of mental development occur in parallel, slumps in creative thinking could occur at the times of jumps in other abilities. Changes in higher-order thinking and reasoning abilities, along with changes in affective processes and emotion-related abilities (understanding, expressing, and regulating emotions), were proposed as especially important for jumps and slumps in creative ability. Lubart and Georgsdottir (2004) proposed that the development of various abilities proceeds in a hierarchical manner in that as one kind of ability is under development, another ability area may be on "hold" or not develop at a similar pace. For example, there may be times when abstract thinking and other analytic abilities are developing, while development of creative abilities is stalled. Lubart and Lautrey (1995) identified a relationship between a slump in creative thinking and an increase in logical reasoning and suggested this is based on a compensating effect.

Creativity in multiple domains includes emotion expression (Botella et al., 2013; Glaveanu et al., 2013), making it important to start understanding how emotion use in performance on creative thinking tests differs across childhood and adolescence. Emotion-related

abilities – understanding and effectively regulating emotions – develop in children and adolescents, showing a generally upward trend, but also some discontinuities and plateaus (Bazhydai, Ivcevic, Widen, & Brackett, 2019; Zimmerman & Iwanski, 2014). Bazhydai and colleagues (2019) found some differences in emotion vocabulary between students in 5th and 6th grade and those in 7th and 8th grades and also some points of stability. Specificity of emotion vocabulary was higher for older students; students in 7th and 8th grade were able to generate more close synonyms for different specific emotions (happy, angry, sad). However, younger and older students generated similar number of total descriptors of different emotion categories. For instance, younger students were likely to include remotely related (but relevant) affective descriptors in different emotion categories (using the term pride when asked for synonyms for happiness). Following suggestions by Lubart and Lautrey (1995), we could hypothesize that use of emotions on creativity tasks would slump at the time of growth in emotion vocabulary; as these adolescents focus on developing emotion-related abilities (more granular, precise, and "correct" emotion vocabulary), creative performance can suffer. Alternatively, greater emotion vocabulary could find its way into creative thinking tasks, especially in the verbal domain.

The asynchronicity in development can interact with the domain specific nature of creativity (Barbot et al., 2016; Ivcevic & Mayer, 2009; Kaufman & Baer, 2005). Many previous studies have used only one type of creativity task, either in the verbal or figural domain, to assess creative ability (e.g., Camp, 1994; Gralewski et al., 2016; Jastrzębska & Limont, 2017; Krampen, 2012). Different creativity domains draw on other abilities (e.g., verbal processing and vocabulary for writing and storytelling creativity, visuospatial ability for figural creativity), which have their own developmental trajectories. These trajectories in turn can create differences in developmental patterns of creative ability across domains, stressing the importance of

studying tasks from multiple domains. A number of recent large-scale studies, crossing cultures (i.e., China, Poland), have administered tasks from only the nonverbal, figural domain, limiting possible conclusions to that domain. The current study sought to examine grade-level differences in creative ability using verbal and figural tasks in a large sample of Spanish students from 1st to 10th grade (primary and secondary school).

Finally, a large proportion of variance in performance on creative thinking tests is task-specific (Barbot et al., 2016; Dumas & Dunbar, 2014). In studies including a range of ages, task-specificity can be related to differential appeal of tasks across ages. Because of this, ideal assessments of creative abilities would administer batteries that include multiple tasks across content domains and modes of creative thinking, such as the EPoC battery (Barbot, Besançon, & Lubart, 2016; Lubart, Barbot, & Besançon, 2019; Lubart, Zenasni, & Barbot, 2013). This approach, while ideal, might not be possible in many instances for practical reasons. We argue that in the cumulative work of science studies that use individual tasks become valuable to test conclusions based on existing studies in the short term and in that they can be aggregated in future meta-analyses representing a broader range of tasks and performance criteria.

## Study Overview

Previous studies on age and/or school grade-related differences in performance on creative thinking tests have not commonly used measures from multiple content domains; of the studies included in the Said-Metwaly et al. (2020) meta-analysis included less than 30% of studies that included both verbal and figural tasks. The current study examined differences in creative ability across a broad range of grade levels, across two task domains, and four performance criteria. Two rarely used tasks were employed – storytelling in the verbal domain and incomplete figure in the non-verbal figural domain. Furthermore, the present study uses

multiple criteria to evaluate performance, including some of the criteria which have been incorporated in a recent meta-analysis (novelty/originality), as well as criteria not commonly used in previous research (elaboration, emotion use). First, we examined slumps and jumps in performance criteria across grade levels, with a focus on the hypothesized slump during transition to secondary school in 7th grade. Second, we compared performance on verbal and figural tasks across grade levels.

## Method

### Sample

A total of 663 students, ranging from 1st-10th grade (ages 5-18 years), attending 5 public school (3 primary and 2 secondary) in Santander, Spain were recruited for the study. The schools included children from a broad range of socio-economic backgrounds from the urban and suburban environments. Students in grades 1-6 attend primary school and students in grades 7-10 attend secondary school. Primary school includes three two-year cycles (1st and 2nd, 3rd and 4th, and 5th and 6th grades) with one primary teacher in each cycle and specialist teachers for music, arts, and physical education. Secondary school includes two two-year cycles (corresponding to 7th and 8th, and 9th and 10th grades), taught by subject matter specialist teachers. Both primary and secondary schools are neighborhood schools, with secondary schools being larger and drawing on more primary schools within the school district.

Although an attempt was made to recruit schools with diverse cultural and socio-economic backgrounds, the sample is best described as one of convenience based on relationships with education researchers at the University of Cantabria and the Botin Foundation.

Sample size by task and grade level can be found in Table 1.

### Measures

*Verbal creativity: Storytelling Task*. Participants were presented with a series of pictures without words from the picture book, "A Boy, a Dog and a Frog" (Mayer, 1967), which tells a short story about a boy and his dog trying to catch a frog in a river. This book has previously been used in creativity research with primary school students (Hoffmann & Russ, 2012, 2016), although it has not been previously used with secondary school students. The task was not timed and students were allowed to spend as much time as they deemed necessary to tell the story.

Because the study included a broad range of student educational levels, task administration had to be adapted to be age appropriate. The youngest students were only learning to read and write and needed assistance to read the instructions and recording their responses. For participants in 1st-3rd grade, the examiner presented the packet of pictures to the child and allowed them to look through the pictures to see what happens in the story. Then the examiner read the following instructions:

> "On these pages are pictures that tell a story about a boy and his dog. Your job is to make up a story that goes along with the pictures. Make it the kind of story you would read in a book. I am going to write down what you say."

The examiner then went through the pictures with the child one at a time and recorded their narrative.

For participants in grades 4-6, the examiner gave these instructions to the entire class of students at once:

> "On the pages in this packet you will find a series of pictures that tell a story about a boy and his dog. Please take a minute to look through all of the pictures so that you know what happens in the story. Then, please make up the story that goes along with the pictures. Make it the kind of story you would read in a book."

The students were then asked to write their responses on a provided testing form.

For grades 7-10, students read the same instructions as those that were read to the participants in grades 4-6, but did so silently, to themselves, before writing out their stories independently, with the examiner serving as the proctor for the class and available to answer any questions.

Two research assistants used the consensual assessment technique (Amabile, 1996) to code children's stories. These coders worked for at least one semester in a creativity research lab and completed readings on definition of creativity as part of their training. Following Amabile (1996), they were provided brief descriptions of the four performance criteria and instructed to use their subjective judgment in assigning their ratings (see Table 2). Coders assigned a score of 1 (not at all) to 5 (very much) for each of the following variables: creativity, novelty/originality, elaboration, and use of emotions. Before coders started to assign ratings to individual responses, they were presented a random subsample of responses across grades to familiarize themselves with the sample and use to calibrate their judgments to the present sample (as opposed to an absolute criterion of creativity).

The coders were two Spanish-speaking research assistants who independently rated all of the stories, and inter-rater reliabilities were calculated for each of the variables. The alpha coefficients were all in the excellent range (creativity: $\alpha = .90$, novelty/originality: $\alpha = .91$, elaboration: $\alpha = .96$, and emotion use: $\alpha = .91$). Ratings were averaged across raters to provide a single score for each variable.

*Figural creativity: The Incomplete Figure Task.* A drawing task was adapted from a subtest of the Torrance Tests of Creative Thinking (TTCT; Torrance, 1966), which have been used with participants from kindergarten into adulthood (Kim, 2011). Participants were

presented with an abstract shape comprised of three lines and given 3 minutes to complete a drawing using the shape in any way they desired. Participants were also asked to provide titles for their drawings.

Instructions for all students were as follows:

> "Please add to the figure below to make a picture of anything you want. Try to tell a complete story with your picture and then give your picture a title. You will have three minutes to draw and come up with a title. There are no right or wrong answers, so you may draw whatever you wish."

Participants in grades 1-3 had instructions read to them by the experimenter in a one-on-one session. Participants in grades 4-6 had instructions read to them, but completed the measure in a group classroom setting. Participants in grades 7-10 independently read the instructions and completed the task in their classroom.

The Incomplete Figure Task was coded using the same categories as the Storytelling Task: creativity, novelty/originality, elaboration, and emotion use. Two coders assigned a score of 1-5 for each of the 4 variables. Titles/captions were translated from Spanish to English before coding. The inter-rater reliabilities were all within the acceptable to excellent range (creativity: $\alpha = .76$, novelty/originality: $\alpha = .80$, elaboration: $\alpha = .69$, and emotion use: $\alpha = .68$). Scores were averaged to produce a single score for each variable.

**Procedure**

Students were recruited to participate in the research study through letters and consent forms sent home to parents for review. If the consent forms were returned to the school signed by parents, the student was eligible to participate. In addition, before beginning any of the assessments, the students were asked to sign a voluntary informed assent form. Children ages 6-

10 checked a box and printed their name and date on the form. For children ages 11 and older, they were asked to check a box and both print and sign their name with the date.

Students completed assessments during school hours. Testing was conducted using a coding system to assure student anonymity and confidentiality and lasted less than thirty minutes per participant.

## Results

The results are organized in three sections: (1) Preliminary analyses examine measurement invariance of creativity measures across three age groups (with age-related variations in study administration); (2) Grade-related differences for verbal and figural creativity; and (3) Comparison of verbal and figural creativity across grade levels. Raw scores for creative performance measures were transformed to z-scores across grades, which ensured that the two domains have the same mean and variance, thus eliminating potential differences of difficulty by domain attributed to raters' standards for each domain. Furthermore, z-scores offer a means to size the mean performance across age groups.

### *Preliminary analyses*

Because creativity measure administration differed depending on the age group, we tested measurement invariance (MI) in the multi-group confirmatory factor analysis (CFA) framework based on guidelines in the literature (Bollen, 1989, Byrne, Shavelson, & Muthén, 1989). The model specified two correlated factors of creative performance – verbal (storytelling task) and figural (incomplete figures task), with four observed indicators comprising each factor – creativity, novelty, elaboration, and emotion use. This procedure was used to ensure that the creativity performance measures applied similarly to the three administration groups, allowing the meaningful comparison across grade levels. Four levels of MI stringency were tested, each

adding a new set of constraints: (1) the congeneric model was fitted simultaneously to the three age groups, with the only constraint of an identical number of factors and factor loading pattern across groups, (2) the weak model set factor loadings to be equal across groups, (3) the strong invariant model further set indicator intercepts to be equal, and (4) the strict invariant model added the equality of residual variance across the three age groups. Model parameters were estimated using AMOS 26 (Arbucke, 2019).

Model fit was assessed based on the most widely used indexes of model fit (Kline, 2010), including the Chi-square test, Chi-square to degree of freedom ratio, the normed fit index (NFI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). MI decisions were based on the $\Delta\chi^2$ significance and $\Delta$CFI (with value less than 0.01 reflecting MI; Byrne, 2010; Cheung & Rensvold, 2002) using the congeneric model as baseline.

Fit indices of the MI testing for the above described two factor model of figural and verbal creativity (congeneric model: $\chi^2[57] = 99.065$, $p < .001$, $\chi 2/df = 1.738$, CFI = .977, RMSEA = .034; weak invariance model: $\chi^2[69] = 109.367$, $p = .001$, $\chi^2/df = 1.585$, CFI = .978, RMSEA = .030; strong invariance model: $\chi^2[75] = 112.480$, $p = .003$, $\chi^2/df = 1.500$, CFI = .980, RMSEA = .028; strict invariance model: $\chi^2[91] = 219.910$, $p < .001$, $\chi^2/df = 2.416$, CFI = .931, RMSEA = .047) indicated that the strong invariance held across age groups ($\Delta\chi^2_{[df]} = 13.415$ [18], $p = .766$, $\Delta$CFI = 0.002), and that the strict invariance did not hold ($\Delta\chi^2_{[df]} = 120.844$ [34], $p < .001$, $\Delta$CFI = .046). Strong invariance is sufficient for group comparisons (e.g., Barbot et al., 2016; Millsap, 2011). Thus, we compare children across three grade groupings that received somewhat different (developmentally-appropriate) test administration.

*Grade-level differences in verbal and figural creativity domains*

In order to assess differences in creative ability across grade level as measured by the Storytelling and Incomplete Figure Tasks, multivariate analyses of variance (MANOVA) were conducted on the four performance criteria variables (creativity, novelty/originality, elaboration, and emotion use). Post hoc tests (Tukey HSD) were conducted to examine how ability varied by grade level.

*Incomplete Figure Task: Creative Ability in the Figural Domain*

There were significant differences in performance on the Incomplete Figure Task by grade level for creativity, $F(9, 643) = 5.63$; $p < .001$; elaboration, $F(9, 643) = 3.17$; $p = .001$, and emotion use, $F(9, 643) = 2.26$; $p = .017$. No significant differences by grade level were evident in novelty/originality, $F(9, 643) = 1.46$; $p = .161$;(see Table 3 and Figure 1).

Follow-up post hoc pair-wise analyses (see Table 4) showed that for creativity, students in 1st grade scored significantly lower than students in 4th through 10th grades. Students in 2nd grade also scored significantly lower than students in 9th and 10th grade. There were no other significant grade level differences for ratings of creativity.

For elaboration, post hoc analyses showed that students in 4th grade scored significantly lower than students in 3rd and 10th grade. Lastly, students in 1st grade scored significantly lower than students in 6th grade on emotion use.

*Storytelling Task: Creative Ability in the Verbal Domain*

Rating scores across all variables on the Storytelling Task significantly differed by grade level; creativity, $F(9, 643) = 7.83$; $p < .001$; novelty, $F(9, 643) = 3.62$; $p < .001$; elaboration, $F(9, 643) = 5.82$; $p < .001$; and emotion use, $F(9, 643) = 4.62$; $p < .001$ (see Table 3 and Figure 2). Table 5 shows results for post hoc pair-wise comparisons of grade levels. For creativity, students in 10th grade scored higher than students in grades 1 through 5, as well as those in 7th grade.

Also, students in 8th and 9th grades scored higher than those in 1st, 3rd, and 4th grades. Overall, creativity is similar across primary school (grades 1-6) and tends to be higher in secondary school grades when compared with primary school.

Few differences were observed for novelty/originality. Students in 10th grade had higher scores than those in 1st and 4th grades and 9th graders had higher scores than 4th graders. Similarly, elaboration was higher in 10th graders than for 1st, 2nd, 4th, 5th, and 7th graders. Students in 9th grade had higher creativity scores than those in 4th and 5th grade, and 8th graders also scored higher than 5th graders. Finally, most observed differences for emotion use were in relation to 5th graders; they had lower scores than secondary school students (7th through 10th grade), and lower scores in comparison to 2nd through 4th graders, pointing to a 5th grade as a slump in emotion use.

*Domain Differences in Creative Ability*

A within-subject repeated measures ANOVA was conducted to examine how performance on each criterion varied by task domain across grade levels (see Table 6). Performance on the two tasks did not differ across grades, with the exception of creativity and emotion use being higher for figural than verbal creativity in 5th grade students.

Furthermore, correlations between performance criteria on the figural and verbal tasks (e.g., novelty on the figural tasks and novelty on the verbal task) show a similar pattern across grade levels (see Table 7). Across grades, performance on the two domains is not significantly correlated, with a few exceptions in 3rd and 8th grade. However, these correlations should be replicated; applying a Bonferroni correction for multiple comparisons, only two correlations remains significant (emotion use in 3rd grade and novelty in 8th grade).

## Discussion

The current study sought to examine differences in creative ability across grade level, performance criteria, and task domain. Similar to the meta-analysis by Said-Metwaly et al. (2020), within task domain, ratings of creativity tended to show a steady increase with grade level, while other performance criteria showed more pronounced slumps and jumps. Furthermore, we did not find evidence for the general 4th grade slump. However, unlike in the results of the meta-analysis (Said-Metwaly et al., 2020), there was also no evidence of a general 7th grade slump at the of transition from primary to secondary school.

For the figural task, creativity ratings were steadily higher from 1st grade to 6th, followed by a slump in 7th grade and further increases throughout secondary school grades. As this is the last year of primary school, the jump could be due to the students' ability to shift their focus from more academically oriented tasks to those centered on creativity and thus perform higher (Lubart & Georgsdottir, 2004). Also, as the oldest in school, 6th graders are higher in social status than younger students, potentially giving them higher social self-esteem (Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991), which can enable them to be more unconventional without concerns for their social standing, especially on the very open-ended task that is corresponds to lay conceptions of creativity (Glaveanu, 2011, 2014).

Creativity performance on the verbal storytelling task also showed an upward trend, but starting in secondary school. Creativity ratings were not significant from grade to grade in primary school, but creativity performance tended to be higher in secondary than in primary school. One possible explanation for this timing is in the development of verbal ability and narrative skill development necessary for storytelling. Future research should assess general storytelling ability – ability to structure a coherent story with a beginning, middle, and end – and examine whether once that ability reaches a peak or plateau, creative storytelling performance

starts increasing. This would be in line with the interpretation based on complementary waves in development of mental abilities suggested by Lubart and Lautrey (1995); in other words, as one ability is a focus of development, another one is pauses or slumps.

Performance on novelty/originality, elaboration, and emotion use showed few significant differences across grade levels for the figural task. There was a slump for elaboration in 4th grade and a bump in emotion use in 6th grade (comparing to the beginning of elementary school). Lack of differences among grade levels is more notable than the differences in cross-grade comparisons. For the verbal storytelling task, novelty/originality showed a dip in 4th grade, elaboration a dip in 4th and 5th grade, and emotion use slump in 5th grade. For both novelty/originality and elaboration, performance before the 4th grade dip does not significantly differ, but the 4th grade dip becomes apparent comparing to high scores in 9th and 10th grade. Emotion use scores showed a slump in 5th grade, with a significant drop from 4th grade and followed by a significantly higher score in 7th through 10th grade. The different developmental patterns for the two creative ability domains suggest that performance might be influenced by development of domain-relevant skills.

It is important to consider the observed differences in emotion use, especially in light of this performance measure not been previously used in research. A conceptually similar variable, humor and affectivity, was used by Gralewski and colleagues (2016) for a figural task, Test for Creativity Thinking – Drawing Production (showing a steady increase from preschool to college). Studies that do not assess emotion-related abilities, such as emotion vocabulary and emotion regulation, are not able to directly address possible reasons for the observed differences across age or grade levels. One hypothesis is that emotion use on verbal creativity tasks is related to the development of emotion vocabulary. Research examining emotion vocabulary in early

adolescents shows that 5th graders produce the least number of distinct synonyms for different emotion categories (Bazhydai et al., 2019). As the emotion vocabulary increases, so could the salience of using emotions in creative storytelling. Another hypothesis is that emotion use on the verbal creativity task is related to development of emotion regulation abilities. Compared with older adolescents, children at age 11 (corresponding approximately to 5th grade) show high use of the adaptive emotion regulation strategies, but also high use of some maladaptive strategies, such as passivity or dysfunctional rumination **(**Zimmerman & Iwanski, 2014**)**. This pattern of emotion regulation strategy use suggests a developmental focus on learning to distinguish effective from ineffective strategies. As development focuses on learning about appropriate emotion regulation, the use of emotions in the service of creative storytelling (potentially including remotely related emotions) might be suppressed, in line with the proposition by Lubart and Lautrey (1995).

Differences in performance across tasks observed in this study coupled with extensive research showing that creativity is at least in part domain specific (Barbot et al., 2016; Han & Marvin, 2002; Ivcevic & Mayer, 2006, 2009; Lubart & Guignard, 2004; Silvia, Kaufman, & Pretz, 2009) attest to the importance of studying the development of creative ability in different domains. Strong evidence for domain specificity of creativity comes from studies of creative performance where participants create multiple products, such as stories, mathematics problems, puzzles, or drawings, and the correlations among the creativity ratings for these products are consistently low (Baer, 1998). Supporting these performance differences, neuroscientific research shows differential brain activation when individuals engage in tasks pertaining to different creativity domains. For instance, Boccia, Piccardi, Palermo, Nori, and Palmiero (2015) found that musical creativity produces activation in the bilateral medial frontal gyrus, the left

cingulate gyrus, and middle frontal gyrus, whereas verbal creativity activates areas mainly in the left hemisphere, prefrontal cortex, and middle and superior temporal gyri. Without using measures that capture multiple domains, findings about development of creative ability may prove incomplete and overgeneralized.

Not only did task domain impact creative ability across grade levels, but different performance criteria also varied in their relationship with creativity in each task domain. For example, in the Storytelling Task, there was a dip in elaboration and emotion use in 5$^{th}$ grade that was not accompanied by a dip in creativity, suggesting that creativity may be achieved by means other than elaboration or emotion use at that grade level. Future research should explicitly examine what performance attributes predict creativity and whether this varies at different ages or across domains.

**Limitations and Future Directions**

Although this study has strengths in regards to sample size, range of grade levels assessed, and assessments in two domains, it is important to note certain limitations. The cross-sectional nature of the study does not allow conclusions about developmental change in the same individuals and introduces potential biases due to cohort and sampling bias effects (Barbot, 2019). Future research would ideally be able to focus on designing studies that allow for the developmental time course of creativity to be delineated through longitudinal and cross-sequential designs. These designs are costly and thus rare. To date, there is only one cross-sequential study in which creativity thinking tests were administered to five cohorts of children (two kindergarten and first through third grades of elementary school) and then followed for three years (Krampen, 2012).

With these limitations noted, cross-sectional designs have some advantages too. In a review of methods for developmental study of creativity, Barbot (2019) argued that cross-sectional designs could describe (noncausal) developmental trends – the goal of the present study – more realistically and meaningfully than longitudinal designs because they avoid the issues associated with repeated measurement and stimulus dependency. Retesting with the same measures introduces biases based on earlier presentation of the task; the task is not novel any more and participants can rely on previously generated responses or try to inhibit these responses. In adults, bias due to retest exists for long periods of time (up to 7 years) for a range of cognitive tasks (Salthouse, Schroeder, & Ferrer, 2004), although it is not clear how long such effects exist for children and adolescents. Using alternate forms of measures in longitudinal studies is also problematic; seemingly alternate forms of the same measures tend to have low alternate-form reliability, suggesting that measurement is dependent (to substantial extent) on the stimulus content.

The contribution of the present study is in the administration of not commonly used measures and scoring of tests using not oft scored performance criteria, such as elaboration and emotion use, which will be useful in future meta-analyses. However, a limitation of the study is that creative ability was assessed with only one task in verbal and figural domains. Barbot and colleagues (2016) provide evidence that a substantial proportion of variance in performance on tests of creative thinking is task-specific. Thus, the observed results might have been affected by the specific nature of these tasks. For instance, it may be more difficult to be novel or original in telling a story that is relatively well defined through a series of pictures (as in the present study) than it is to complete a figure that does not present with any preset form. Emotions may be more readily expressed verbally than in figural creativity. To address this, researchers should

administer multi-domain test batteries. One example of such battery is the EPoC (Barbot et al., 2016; Lubart et al. , 2019; Lubart et al., 2013). This battery includes measures of both divergent-exploratory and convergent-integrative creative thinking processes, as well as assessing performance across content domains. With multiple items in visual art, verbal-literary, social, scientific, math, music, and body movement domains, the EPoC battery can increase reliability of conclusions about developmental trends across domains.

Of note, the sample in the present study was not a representative one randomly sampled from the population of primary and secondary school students. However, the sample included students of all ability levels, it was from urban and suburban areas, and drew from a broad range of socio-economic backgrounds. Furthermore, the gender composition of the sample was not uniform across grades. Previous research has largely shown a lack of mean gender differences in creative abilities, suggesting a "relative equality" in creativity development among males and females (Baer & Kaufman, 2008). However, a more recent line of research examined gender differences in variability of performance scores (some finding support for greater male variability: He, 2018; He & Wong, 2011; Karwowski, Jankowska, Gralewski, Gajda, Wiśniewska, Lebuda, 2016, and some not finding differences: Taylor & Barbot, 2021), stressing the need for future research to examine both age/grade level and gender differences.

Important to note is that the presentation of figural and verbal tasks was not counterbalanced and that the method of task administration varied across grades; students in the youngest grades (1-3) told their stories verbally, those in the middle grade group (grades 4-6) were read the instructions and wrote their answers, and those in the oldest grades (grades 7-10) read the instructions and wrote their responses independently. The tasks in lowest grades were also administered in a one-on-one setting versus group administration in later grades. The

variations in administration were guided by the constraints of children's age and developmental and educational level. For instance, as children are just learning how to read and write at the start of elementary school, we decided to administer the measures one-on-one in 1st to 3rd grade. Future work should use technology to enable even the youngest students to independently complete the tasks (e.g., administer tasks on mobile devices that can record either spoken or written narratives). To test whether performance across grade levels associated with these different administration procedures behaved similarly, we tested measurement invariance of the creativity measures. Evidence of strong invariance showed that it was warranted to compare these three grade level groups on their creativity performance.

Further, relevant measures of participant functioning outside creative ability – cognitive measures of intelligence or vocabulary and socioemotional development variables, such as understanding emotions – were not obtained as part of the current study. This limits the ability to understand what portion of difference in performance by grade level may be due to changes in other domains of development (as suggested by Lubart & Georgsdottir, 2004). In addition to development of abstract and logical thinking abilities in adolescence that have been discussed by Lubart and Lautrey (1995), recent research points to the importance of creative meta-cognition abilities for creativity (Anderson & Haney, 2020). These are abilities to reflect on the creative process to explicitly identify strategies used and potential ways of improving performance. Assessing creative metacognition in future research will make it possible to address to what extent children and (especially) adolescents are using different approaches to the task in the service of creativity (e.g., elaboration or expressing emotions). Finally, the current study also assessed creative ability within a relatively culturally homogenous sample of Spanish students. Future work in comparing trajectories of development across verbal and figural task domains

could benefit from cross-cultural comparisons, especially including countries with different educational systems (timing of school trajectories, educational tracking).

**Conclusion**

The current study examined grade differences in creative ability across verbal and figural creativity tasks with a large sample size and age range. Findings provide evidence that creative ability on multiple performance criteria significantly vary by grade level. Task domain made a difference in performance as well. More jumps and slumps across grades were evident in the verbal Storytelling Task. In sum, results point to a complex relationship between grade level, task domain, and creativity-related performance criteria, and the need for more research across these areas. Given the effects of creativity across functioning, refining our understanding of the developmental course of creative ability over the school-age years is important for accurate assessment and possible intervention to develop creative ability.

Acknowledgement

The work on this paper was funded by and done in collaboration with the Botin Foundation (*Emotions, Creativity and the Arts* grant; Principal investigator: Zorana Ivcevic).

References

Adler, P. A., & Adler, P. (1995). Dynamics of inclusion and exclusion in preadolescent cliques. *Social Psychology Quarterly, 58*, 145–162.

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.

Amabile, T. M. (1996). *Creativity in context*. Westview.

Andrew, M., & Flashman, J. (2017). School transitions, peer influence, and educational expectation formation: Girls and boys. *Social Science Research*, *61*, 218-233.

Arbuckle, J. L. (2019). Amos 27.0 User's Guide. IBM SPSS.

Baer, J. (1998). The case for domain specificity of creativity. *Creativity Research Journal*, *11*(2), 173-177.

Baer, J., & Kaufman, J. C. (2008). Gender differences in creativity. *Journal of Creative Behavior, 42*(2), 75-105. https://doi.org/10.1002/j.2162-6057.2008.tb01289.x

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. (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

Barbot, B., Hein, S., Luthar, S. S., & Grigorenko, E. L. (2014). Capturing age-group differences and developmental change with the BASC Parent Rating Scales. *Journal of Applied Developmental Psychology, 35*(4), 294–303. https://doi.org/10.1016/j.appdev.2014.04.003

Barbot, B., Lubart, T. I. & Besançon, M. (2016). “Peaks, slumps, and bumps”: Individual differences in the development of creativity in children and adolescents. In B. Barbot (Ed.), Perspectives on creativity development. *New Directions for Child and Adolescent Development, 151*, 33–45.

Bazhydai, M., Ivcevic, Z., Widen, S., & Brackett, M. A. (2019). Emotion vocabulary in early adolescents. *Imagination, Cognition, and Personality, 38(4),* 378–404. doi: 10.1177/0276236618765403

Boccia, M., Piccardi, L., Palermo, L., Nori, R., & Palmiero, M. (2015). Where do bright ideas occur in our brain? Meta-analytic evidence from neuroimaging studies of domain-specific creativity. *Frontiers in Psychology, 6*, Article 1195. Doi: 10.3389/fpsyg.2015.01195.

Bollen, K. A. (1989). *Structural Equations with Latent Variables. Wiley Series in Probability and Mathematical Statistics*. Wiley.

Botella, M., Glaveanu, V., Zenasni, F., Storme, M., Myszkowski, N., Wolff, M., & Lubart, T. (2013). How artists create: Creative process and multivariate factors. *Learning and Individual Differences, 26*, 161–170. https://doi.org/10.1016/j.lindif.2013.02.008

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Camp, G. C. (1994). A longitudinal study of correlates of creativity. *Creativity Research Journal, 7*, 125-144.

Charles, R. E., & Runco, M. A. (2000-2001). Developmental trends in the evaluative and divergent thinking of children. *Creativity Research Journal, 13(3-4),* 417-437.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255.

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

Cropley, A. J. (2001). *Creativity in education & learning: A guide for teachers and educators*. Psychology Press.

Dahlman, S., Bäckström, P., Bohlin, G., & Frans, Ö. (2013). Cognitive abilities of street children: Low-SES Bolivian boys with and without experience of living in the street. *Child Neuropsychology*, *19*(5), 540-556.

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.

Daugherty, M. (1993). Creativity and private speech: Developmental trends. *Creativity Research Journal*, *6*(3), 287-296.

Dede, C. (2010). Comparing frameworks for 21st century skills. In J. Bellanca & R. Brandt (Eds.), *21st century skills* (pp. 51–76). Solution Tree Press.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (p. 466–503). Oxford University Press. Doi: 10.1093/acprof:oso/9780195134971.003.0029.

Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychonomic bulletin & review*, *11*(6), 1011-1026.

Dumas, D., & Dunbar, K. N. (2014). Understanding fluency and originality: A latent variable perspective. *Thinking Skills and Creativity, 14*, 56–67. doi:10.1016/j.tsc.2014 .09.003

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1997). Development during adolescence: The impact of stage–environment fit on young adolescents' experiences in schools and in families (1993). In J. M. Notterman (Ed.), *The evolution of psychology: Fifty years of the American Psychologist* (p. 475–501). American Psychological Association. Doi: 10.1037/10254-034.

Forgeard, M. J. C., & Kaufman, J. C. (2016). Who cares about imagination, creativity, and innovation, and why? A review. *Psychology of Aesthetics, Creativity, and the Arts, 10*(3), 250–269. https://doi.org/10.1037/aca0000042

Gajda, A., Karwowski, M., & Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology*, *109*(2), 269.

Glaveanu, V. P. (2011). Is the lightbulb still on? Social representations of creativity in a western context. *International Journal of Creativity and Problem Solving, 21*, 53–72. Retrieved from http://eprints.lse.ac.uk/39032/.

Glaveanu, V. P. (2014). Revisiting the ‘‘art bias’’ in lay conceptions of creativity. *Creativity Research Journal, 26*, 11–20. doi:10.1080/10400419.2014.873656.

Glăveanu, V. P., Lubart, T., Bonnardel, N., Botella, M., de Biaisi, P.-M., Desainte-Catherine, M., Georgsdottir, . . . & Zenasni, F. (2013). Creativity as action: Findings from five creative domains. *Frontiers in Psychology, 4*(176), 1–14. doi: 10.3389/ fpsyg.2013.00176

Gocłowska, M. A., Damian, R. I., & Mor, S. (2018). The diversifying experience model: Taking a broader conceptual view of the multiculturalism–creativity link. *Journal of Cross-Cultural Psychology, 49*(2), 303–322. https://doi.org/10.1177/0022022116650258

Gralewski, J., Gajda, A., Wisniewska, 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. DOI: https://doi.org/10.1515/ctra-2016-0011

Han, K. S., & Marvin, C. (2002). Multiple creativities? Investigating domain-specificity of creativity in young children. *Gifted Child Quarterly*, *46*(2), 98-109.

He, W. (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*, 1-14, doi: 10.3389/fpsyg.2018.02331

He, W., & Wong, W. (2011). Gender differences in creative thinking revisited: Findings from analysis of variability. *Personality and Individual Differences, 51*(7), 807-811. Doi: 10.1016/j.paid.2011.06.027

Hoffmann, J., & Russ, S. (2012). Pretend play, creativity, and emotion regulation in children. *Psychology of Aesthetics, Creativity, and the Arts, 6*(2), 175–184. https://doi.org/10.1037/a0026299

Hoffmann, J. D., & Russ. S. (2016). Fostering pretend play skills and creativity in elementary school girls: A group play intervention. *Psychology of Aesthetics, Creativity, and the Arts, 10*(1), 114-125. https://doi.org/10.1037/aca0000039

Isen, A. M. (1999). Positive affect and creativity. In S. W. Russ (Ed.), *Affect, creative experience, and psychological adjustment* (pp. 3–17). Brunel/Mazel.

Isen, A. M., Johnson, M. M., Mertz, E., & Robinson, G. F. (1985). The incluence of positive affect on the unusualness of word associations. *Journal of Personality and Social Psychology, 48*, 1413–1426.

Ivcevic, Z., & Mayer, J. D. (2006). Creative types and personality. *Imagination, Cognition and Personality*, *26*(1), 65-86.

Ivcevic, Z., & Mayer, J. D. (2009). Mapping dimensions of creativity in the life-space. *Creativity Research Journal*, *21*(2-3), 152-165.

Jastrzębska, 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*(3), 337–342. https://doi.org/10.1080/10400419.2017.1360060.

Karwowski, M., Jankowska, D. M., Gralewski, J., Gajda, A., Wiśniewska, E., Lebuda, I. (2016). Greater male variability in creativity: A latent variables approach. *Thinking Skills and Creativity, 22*, 159-166, doi: 10.1016/j.tsc.2016.10.005

Kaufman, J. C., & Baer, J. (Eds.). (2005). *Creativity across domains: Faces of the muse*. Psychology Press.

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.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Urban, K. K. (1991) On the development of creativity in children, *Creativity Research Journal, 4*(2), 177-191, doi: 10.1080/10400419109534384

Kleibeuker, S. W., De Dreu, C. K. W., & Crone, E. A. (2013). The development of creative cognition across adolescence: Distinct trajectories for insight and divergent thinking. *Developmental Science, 16(1),* 2-12.

Krampen, G. (2012). Cross-sequential results on creativity development in childhood within two different school systems: Divergent performances in Luxembourg versus German kindergarten and elementary school students. *Europe's Journal of Psychology, 8*, 423-448, doi:10.5964/ejop.v8i3.468

Lau, S., & Cheung, P. C. (2010). Developmental trends of creativity: What twists and turns do boys and girls take at different grades? *Creativity Research Journal, 22,* 329-336. doi:10.1080/10400419.2010.503543

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.

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.

Lubart, T. I., & Lautrey, J. (1995). *Relationships between creative development and cognitive development.* Paper presented at the Seventh European Conference on Developmental Psychology, Krakow, Poland.

Lubart, T. I., Zenasni, F., & Barbot, B. (2013). Creative potential and its measurement. *International Journal of Talent Development and Creativity, 1*(2), 41–51.

Mayer, M. (1967). *A Boy, a Dog, and a Frog*. Dial Press.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Routledge/Taylor & Francis Group.

Pelowski, M., Specker, E., Gerger, G., Leder, H., & Weingarden, L. S. (2020). Do you feel like I do? A study of spontaneous and deliberate emotion sharing and understanding between artists and perceivers of installation art. *Psychology of Aesthetics, Creativity, and the Arts, 14*(3), 276–293. https://doi.org/10.1037/aca0000201

Runco, M. A., & Charles, R. E. (1997). Developmental trends in creative potential and creative performance. In R. A. Runco (Ed.), *The creativity research handbook* (pp. 115-152). Hampton Press.

Runco, M. A., Pritzker, M. A., Pritzker, S. R., & Pritzker, S. (Eds.). (1999). *Encyclopedia of creativity* (Vol. 2). Elsevier.

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*. https://doi.org/10.1007/s10648-020-09547-9

Salthouse, T. A., Schroeder, D. H., & Ferrer, E. (2004). Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. Developmental Psychology, 40, 813–822. http://dx.doi.org/10.1037/0012-1649.40.5.813

Silvia, P. J., Kaufman, J. C., & Pretz, J. E. (2009). Is creativity domain-specific? Latent class models of creative accomplishments and creative self-descriptions. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(3), 139.

Smith, G. J., & Carlsson, I. (1983). Creativity in early and middle school years. *International Journal of Behavioral Development, 6,* 167-195.

Smith, G. J., & Carlsson, I. (1985). Creativity in middle and late school years. *International Journal of Behavioral Development,8,* 329–343.

Smith, G. J., & Carlsson, I. M. (1990). The creative process: A functional model based on empirical studies from early childhood to middle age. *Psychological Issues*, *57*, 1-243.

Taylor, C. L., & Barbot, B. (2021). Gender differences in creativity: Examining the greater male variability hypothesis in different domains and tasks. *Personality and Individual Differences.* Online first publication. https://doi.org/10.1016/j.paid.2021.110661.

Torrance, E. (1963). *Education and the creative potential*. University of Minnesota Press.

Torrance , E. P. (1966). *The Torrance Tests of Creative Thinking– Norms: Technical Manual Research Edition—Verbal Tests, Forms A and B— Figural Tests, Forms A and B* . Personnel Press .

Torrance, E. P. (1967). *Understanding the fourth grade slump in creative thinking* (Report No. BR-5-0508; CRP-994). U. S. Office of Education.

Torrance, E. P. (1968). A longitudinal examination of the fourth-grade slump in creativity. *The Gifted Child Quarterly, 12*(4), 195–199. https://doi.org/10.1177/001698626801200401.

Urban, K. K. (1991). On the development of creativity in children, *Creativity Research Journal, 4(2)*, 177-191. doi: 10.1080/10400419109534384

Wigfield, A., Eccles, J. S., Mac Iver, D., Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental psychology*, *27*(4), 552.

World Economic Forum (2020). The future of jobs report. World Economic Forum. Retrieved from http://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf.

Zimmermann, P., & Iwanski, A. (2014). Emotion regulation from early adolescence to emerging adulthood and middle adulthood: Age differences, gender differences, and emotion-specific developmental variations. *International Journal of Behavioral Development, 38*(2), 182–194. https://doi.org/10.1177/0165025413515405