Thinking Skills and Creativity 40 (2021) 100814

Contents lists available at ScienceDirect

Thinking Skills and Creativity

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

# Divergent thinking in four-year-old children: An analysis of thinking processes in performing the Alternative Uses Task

Honghong Bai (白红红)[a,*], Hanna Mulder[a], Mirjam Moerbeek[b], Evelyn H. Kroesbergen[c], Paul P.M. Leseman[a]

[a] *Department of Pedagogy and Education: Development & Education of Youth in Diverse Societies, Utrecht University, Heidelberglaan 1, 3584 CS, Utrecht, the Netherlands*
[b] *Department of Methodology and Statistics, Utrecht University, Padualaan 14, 3584 CH, Utrecht, the Netherlands*
[c] *Behavioural Science Institute, Radboud University, PO Box 9104, 6500 HE, Nijmegen, the Netherlands*

A R T I C L E I N F O

*Keywords:*
Divergent thinking
Executive processes
Isolating and recombining properties
Young children
Alternative uses task
4-Year-old

A B S T R A C T

The present–> study examined the divergent thinking (DT) processes of four-year-old children, as part of a longitudinal project that investigates the development of DT in children. Following a similar approach used in a study with adults, children were encouraged to report on their thinking processes through interactive dialogues while performing a widely used DT task, the Alternative Uses Task (AUT). Content analysis of children's utterances revealed that children generated uses mostly based on automatic, bottom-up associative processes and occasionally based on effortful, top-down executive processes. Using (multilevel) regression analysis, we found that (1) both associative and executive DT processes predicted children's fluency scores on the AUT, whilst only the executive DT process *Performing mental operations on the stimulus* uniquely predicted originality; (2) children at the age of four years already showed a serial order effect in the originality of their responses, indicating that the originality of uses increased the later a particular use was generated in the series of mentioned uses; and (3) similar serial order effects characterized the occurrence of executive processes. These results suggest that increasing originality depends on increasing involvement of effortful executive processes. Especially the executive process of mentally isolating properties or parts of objects and the subsequent recombination of these parts and properties into a new structured whole might be a key characteristic of DT to generate original ideas.

## 1. Introduction

Since the 1950s, the study of creativity has been a topic of great interest in various disciplines, including psychology, education, and cognitive and neurobiological sciences (Guilford, 1950; Silvia, 2017). The recent surge in creativity studies, including studies on divergent thinking as a main component of creativity, relates to the increased importance attached to creativity and divergent thinking as typical 21st century skills (Ananiadou & Claro, 2009). Divergent thinking (DT) is defined as a thought process or thinking method used towards exploring multiple solutions (Wang, Hao, Ku, Grabner, & Fink, 2017) and is of particular importance for creativity as it

* Corresponding author.
*E-mail addresses:* honghong.bai@outlook.com (H. Bai), h.mulder2@uu.nl (H. Mulder), m.moerbeek@uu.nl (M. Moerbeek), e.kroesbergen@pwo.ru.nl (E.H. Kroesbergen), p.p.m.leseman@uu.nl (P.P.M. Leseman).

https://doi.org/10.1016/j.tsc.2021.100814
Received 14 February 2020; Received in revised form 17 November 2020; Accepted 9 March 2021
Available online 17 March 2021

enables individuals to think in multiple directions (Charles & Runco, 2001; Kuhn & Holling, 2009). Research on DT to date has shown that even one- and two-year-old children are already able to think divergently (Bijvoet-van den Berg & Hoicka, 2014; Hoicka et al., 2016). Studies on somewhat older children have focused mainly on describing the developmental level of DT at a certain age or over time (Barbot, Lubart, & Besançon, 2016; Gralewski, Lebuda, Gajda, Jankowska, & Wiśniewska, 2017; Krampen, 2012; Torrance, 1968). As such, the thinking processes underlying novel idea generation on DT tasks in children is currently not well understood. Such knowledge is important, as it may contribute to the development of activities and guide interactions to foster DT in children, for example in education. The present study addresses this issue through an extensive analysis of four-year-olds' thinking when performing the Alternative Uses Task (AUT; Guilford, 1967)—a commonly used DT task. First, we investigated the process of generating original ideas on the AUT in four-year-olds. To this end, we adapted an established coding scheme that was originally designed for analyzing the process of DT in adults (Gilhooly, Fioratou, Anthony, & Wynn, 2007). Second, we explored the relationship between the processes underlying idea generation and fluency and originality as performance measures of DT. Third, we examined how the timing of the process of idea generation is related to the timing of the generation of original ideas, through investigating the so-called serial order effect of DT (for the first time reported by Christensen, Guilford, & Wilson, 1957).

### *1.1. Research into the process of DT in adults*

In the AUT, participants are requested to generate as many original responses for a common object (e.g., a brick) as possible. Previous studies have shown that responses on such tasks typically follow a clear pattern in which participants initially generate many relatively mundane ideas, followed by increasingly more original responses but generated at a slower pace. This response pattern has been observed consistently from childhood, in children as young as seven years of age (Ward, 1969) to adulthood (e.g., Gilhooly et al., 2007; Wang et al., 2017), and has been labeled 'the serial order effect' of DT (Beaty & Silvia, 2012; Christensen et al., 1957).

Studies among adults into the process of DT have converged on the idea that two types of processes—associative processes and executive processes—are simultaneously involved in novel idea generation on DT tasks and underlie the occurrence of the serial order effect (e.g., Barr, Pennycook, Stolz, & Fugelsang, 2015; Beaty, Silvia, Nusbaum, Jauk, & Benedek, 2014; Hass, 2017a; Sowden, Pringle, & Gabora, 2015). Associative processes, also referred to as "Type 1" processes, are thought to be bottom-up processes which occur rapidly, unconsciously, automatically, and cost little effort (Sowden et al., 2015). Empirical studies have suggested that, when subjects are presented with common objects and requested to generate as many different and creative uses as possible, their semantic memory in relation to these objects is automatically activated (Benedek, Könen, & Neubauer, 2012; Hass, 2016, 2017a, 2017b; Mednick, 1962). Associative theory (Mednick, 1962) postulates that participants often begin with generating mundane uses because mainly semantic elements that are strongly related to the given prompt are activated in the early phase of idea generation. Only after a while, when the most obvious and most strongly associated semantic elements are exhaustively used (with a relatively small proportion of really novel uses), semantic elements that are peripherally related to the given prompt become activated and the proportion of less conventional, really novel uses will rise.

Executive processes, also referred to as "Type 2" processes, refer to a class of top-down processes that influence or constitute divergent thinking, including inhibition, shifting, and working memory, the use of particular (metacognitive) thinking strategies, and other forms of controlled cognitive processes (Beaty et al., 2014; Gilhooly et al., 2007; Sowden et al., 2015). In contrast to the associative processes, executive processes are thought to be relatively slow, top-down controlled, and effortful in nature and, therefore, also more accessible for conscious reflection (Sowden et al., 2015). Recent studies have found that measures of executive functions (Lee & Therriault, 2013; Sharma & Babu, 2017) and other types of effortful thinking as implicated in fluid intelligence (Beaty et al., 2014; Forthmann, Wilken, Doebler, & Holling, 2019; Gilhooly et al., 2007; Krumm, Arán Filippetti, & Gutierrez, 2018; Nusbaum & Silvia, 2011) are closely associated with generating original ideas on the AUT and other DT tasks. According to Beaty and Silvia (2012), the serial order effect in the originality of generated ideas, as discussed above, could possibly be explained by an increasing role of executive processes in DT over time (note, however, that research to date shows inconsistent findings regarding the direction of the association between executive function and DT, e.g., Radel, Davranche, Fournier, & Dietrich, 2015).

In a study among adults investigating the process of DT, Gilhooly et al. (2007) specifically addressed the associative and executive processes that underlie performance in DT. A group of university students were asked to think aloud while performing the AUT. Afterwards, their responses were analyzed with respect to the type of thinking processes that were involved in the generation of ideas. Three main findings stand out. First, the students were often not able to express the thinking processes involved in generating uses, probably because these processes were mostly quick, automatic, bottom-up, and effortless and, therefore, not easily accessible for conscious reflection. Second, if the students could explain their thinking processes, the vast majority of responses was indicative of the involvement of memory-based associative processes in generating alternative uses (cf. Hass, 2017b). Third, although pertaining to only a small minority of the generated uses, students' utterances were also indicative of more controlled and effortful executive processes. For example, participants mentioned that they imagined disassembling an object and use a part of it to generate a novel use (e.g., "remove the laces from the shoe and use them to tie your hair up"). Thinking processes like these go beyond automatic activation of well-entrenched semantic networks and episodic representations in long-term memory, and require additional attentional effort, selective encoding, mental operations such as rotation, stimulus enlargement or reduction, or movement simulation in working memory, alongside with top-down control. Interestingly, and in line with the hypotheses regarding Type 1 and Type 2 processes discussed above, the results of Gilhooly et al. (2007) confirmed that the thinking processes underlying DT performance are differentially related to the main overall outcomes of the AUT. The frequency with which the adult subjects referred to associative memory-based processes when thinking aloud predicted the overall fluency and originality scores, while the frequency of references made to top-down executive processes specifically predicted the originality scores.

### 1.2. The present study

Previous research has suggested that both associative and executive processes are involved in DT, and that they may differentially influence the quantity (i.e., fluency) and quality (i.e., originality) of generated ideas. However, this conclusion is based on studies with adults who have had many years of formal schooling and ample life experiences, leaving unclear if similar thinking processes underlie DT in much younger subjects. Therefore, the aim of the current study was to investigate the process of DT in four-year-olds, in order to elucidate (1) which thinking processes are involved in DT in this young age group, (2) how DT processes relate to DT ability as indicated by fluency and originality, and (3) whether the serial order effect of DT responses—a key phenomenon of DT observed in older children and adults—is already present at this age, and how DT processes may relate to this effect.

Following a methodology similar to Gilhooly et al. (2007), the present study assessed children's thinking processes through their oral reports on the AUT task. The AUT was administered in interactive dialogues as recommended by van Someren, Bernard, and Sandberg (1994); for an example of measuring narrative abilities in young children using prompted story [re]telling, see Scheele, Leseman, Mayo, and Elbers (2012), so that children were actively prompted to give more ideas and to explain their thinking processes during the AUT test. Young children, age four or five years, are not yet well capable of unsupported introspection into their mental processes (Flavell, Green, & Flavell, 2000), which is partly due to their stage of language development. If, however, children of this age are stimulated by cues or involved in supportive dialogues, they reveal unexpected insights in their own cognition (Mercer & Littleton, 2007). A further advantage of this method is that it creates a relatively natural test situation for young children.

Based on the findings of Gilhooly et al. (2007), while considering the general level of children's cognitive and executive function development at four years of age, we hypothesized that: (1) Children would mostly generate uses based on the automatic associative activation of semantic and episodic long-term memory when performing the AUT, resulting in a relatively high frequency of what we coded as memory-based processes (Hypothesis 1); (2) The involvement of top-down executive processes would still be limited, resulting in relatively low frequencies of references made by the children to what we coded as executive processes (Hypothesis 2); (3) Applying memory-based processes would be predictive of the number of uses that children would generate (i.e., fluency) but might not be predictive of the originality of uses, whereas applying executive processes, if any, would be predictive of the originality of uses but might not be predictive of fluency (Hypothesis 3); (4) Children's responses would show a serial order effect, that is, generated uses would become increasingly original the later in the series of generated uses (Hypothesis 4); and (5) Similarly, there would be a serial order effect in references made by the children to the involvement of executive processes in generating uses during the AUT, more specifically, references to top-down controlled thinking and executive processes would occur more frequently the later in the series (Hypothesis 5).

## 2. Method

### 2.1. Participants

The current study is part of a longitudinal research project that investigates the development of DT in children from age four to age six years. Participants were recruited from the kindergartens of four typical primary schools in the Netherlands. These schools were located in neighborhoods which consisted of mostly middle-class SES families. Two schools were approached via personal networks. The other two schools were recruited through snowball sampling via the principal of one of the aforementioned schools. All schools had shown interest in the topic of creativity in education. All four schools provided regular education to students, and no pre-selection of students was conducted in any form. Kindergarten in the Netherlands is part of the primary school system and comprises the first two of a total of eight grades, in which a play-based child-centred curriculum is implemented in mixed-age groups. Initial instruction in academic skills is usually postponed to the last half year before the transition to third grade at age six years, to prepare children for formal education. The description of the longitudinal project and the advertisement for recruiting participants were sent to parents via schools. Parents registered their children for taking part in the project and provided active parental approval in written form or via email. Eventually, hundred seven children ranging from 3.87 to 5.10 years of age (49 boys and 58 girls; age: $M = 4.44$ years, $SD = 0.26$) were enrolled in the project. Based on available data, we estimate the positive response rate to be about 70%. Data used in the current study were collected during the first measurement wave of the longitudinal project and children were in first kindergarten grade at that time. The project was approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences of Utrecht University in 2016 (reference number: FETC16-066).

### 2.2. Measures

The Alternative Uses Task (AUT) was used to measure children's DT ability as well as the thinking processes underlying DT. Pictures of six common objects (size A4) were used as test stimuli, including a hand towel, a brick, a fishnet, a basket, a broom, and a spoon. Before the start of the test, the experimenter explained to children that a number of pictures of common objects would be presented to them and that they would be requested to think of as many different unusual uses as possible of the objects. Subsequently, children were shown a real newspaper. The experimenter first asked what children thought they could do with a newspaper, to start the conversation. Next, the experimenter gave three examples of unusual uses for the newspaper and explained how she came up with these uses, including demonstrations of both associative (e.g., "if you are going to paint, I have seen once that my mother used a newspaper to keep the table clean") and executive thinking processes (e.g., "you can fold the newspaper, then you have a hat"; the experimenter folded a hat while explaining). Occasionally additional examples were given if children did not fully understand the instruction. Next, children were encouraged to generate other unusual uses and explain their thinking processes. When the experimenter thought that children had understood the purpose of the task, she presented the test stimuli, one by one in a randomized sequence, and asked children to generate as many different and unusual uses as possible for these stimuli.

In contrast to studies with adults (e.g., Gilhooly et al., 2007) in which subjects only receive instructions before the actual test begins, we embedded the six AUT stimuli in interactive dialogues with the children. The experimenters actively prompted the children in three ways: (1) to think of more ideas (referred to as 'more-idea prompts' from here on), e.g., "What else can you use a basket for?", (2) to elaborate on their ideas ('explain-idea prompts'), e.g., "How do you do that?", and (3) to explain their thinking processes ('thinking-process prompts'), e.g., "How did you come up with this idea?". Prompting was adapted to the children. All children were prompted, but children who were less expressive or who gave unclear or sophisticated but difficult to grasp ideas received more prompts. Given the young age of the children, we did not set a strict time limit for the test, in order to allow children sufficient time to express their ideas. In practice, the test took about 15–35 min (including the pretest instruction and occasional breaks). All test sessions were video-recorded.

### 2.2.1. Measures of DT ability

Two measures of DT ability, *fluency* and *originality*, were derived from the AUT, based on the uses that children generated (see Fig. 1 for a visual presentation of the data structure). Fluency reflects the total number of distinct uses generated for a given stimulus. Two uses would be considered as distinct if either the involved actions or the objects that afforded the actions differed. For example, "using a brick to build a house" and "using a brick to build a bridge" were considered as distinct uses. The fluency scores of all stimuli were averaged for the analyses.

For scoring originality, the generated uses were further categorized based on the implicated type of action, in line with past research (e.g., Krampen, 2012; van de Kamp, Admiraal, & Rijlaarsdam, 2016). More specifically, "using a brick to build a house" and "using a brick to build a bridge" were categorized as the same *type* of action, that is, "using a brick to build something" (see Appendix A for a list of action categories used for classifying generated uses). Each distinct type of action implicated in a generated use of a particular stimulus was then given an originality score based on how often this type of action was generated across all participants per stimulus (Hao, Wu, Runco, & Pina, 2015; Kirk & Lewis, 2017) and was calculated as follows: Originality of an action = 1 – (The number of participants who generated this type of action / The total number of participants). Different to fluency, originality scores, thus, were calculated at the level of generated uses and, therefore, the data had a hierarchical structure: the originality scores of the generated uses were nested within stimuli and the stimuli were nested within children (see Fig. 1). We used these data in two different ways in the analyses. First, the originality scores obtained for all types of actions were summed per stimulus and then a mean originality score across stimuli was computed for each child to be used in the first series of regression analyses with data aggregated to the child level. Next, the originality scores nested within stimuli and within children were used in the second series of multilevel regression analyses to examine serial order effects (to be further explained below).

### 2.2.2. Measures of DT processes

In order to analyze children's thinking processes we used the coding scheme of Gilhooly et al. (2007) and adapted it for use with the current age group, based on a pilot with seven four- to six-year-old children who did not participate in the main study. The adaptation followed both a top-down theory-driven and a bottom-up thematic coding approach (Braun & Clarke, 2006): (1) The first author coded

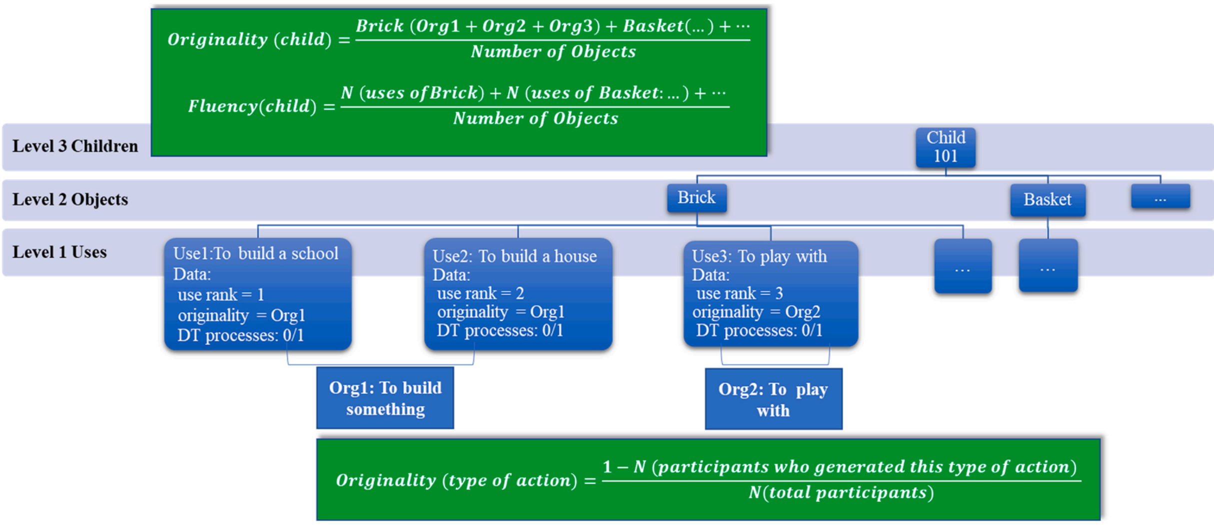


**Fig. 1.** The derivation of fluency and originality scores from the Alternative Uses Task and the nested data structure of originality and divergent thinking processes. The data at level three (child-level) were used in the hierarchical regression analyses, and the data at level one (use-level) were used in the multilevel logistic regression analyses.

the explanations children gave of their thinking processes using Gilhooly et al.'s coding scheme. Explanations that could not be coded were identified and new coding categories were defined based on content analysis of these explanations. (2) A second coder independently applied the updated coding scheme to three transcripts. The coding results of the first and the second coder were compared and discussed, leading to further improvements of the coding scheme. (3) All authors of the present paper reviewed the results of the pilot coding and reached consensus about the categories of Gilhooly et al.'s coding scheme that should be dropped or merged and the new coding categories that should be included. The final coding scheme consisted of six categories to code children's thinking processes (see Table 1).

Coding in the main study proceeded as follows. First, all videos of the AUT test sessions were verbatim transcribed. To facilitate understanding and correct interpretation of children's utterances, the transcripts also included context information related to the test situation, such as children's use of signs and gestures during the test, objects and other elements in the test environment that children referred to, or whether the tests were interrupted (e.g., someone entered the room). Next, the transcripts were divided into episodes, with an episode defined as a coherent stretch of discourse pertaining to a single generated use of a particular stimulus. Then, all episodes were coded for the *presence*(scored 1) or *absence*(scored 0) of each of the thinking processes. Note that the coding categories were not mutually exclusive, and more than one thinking process could occur during a single episode. For example, a child may notice an object in the environment (coded as 2e: 'Mentioning elements in the environment'), that subsequently activates prior knowledge, resulting in new associations based on memory retrieval that lead to the discovery of a new use (coded as 1: 'Retrieval or recall of prior knowledge or experience'). Similar to the data for originality, DT process data were used in the main analyses in two ways. First, for each process category, a child-level average was computed (as the sum of the scores in a particular category across stimuli divided by

**Table 1**
Coding Scheme for Analyzing Children's Explanations of the Divergent Thinking (DT) Processes in the Alternative Use Task, Adapted from Gilhooly et al. (2007).

| DT Process category | Definition | Examples (C = child; T = experimenter) |
|---|---|---|
| 1. Retrieval or recall of prior knowledge or experience | There is clear evidence that the child refers to prior knowledge or prior experience while generating a use. The child may recall a specific memory of a real personal experience or a memory related to here—say, a story, film, or book that relates to the use.<br>Or: The child gives an affirmative answer when asked by the experimenter if he or she had prior personal experience with the use (i.e., if he or she did/learned it before) or if the child knew about the use from others, stories, movies, or books. | 1. C: "I always do that with my father."<br>2a. T: "Have you done that before?"<br>2b. C: "Yes, I have done it once." |
| 2a. Performing mental operations on the stimulus | The child mentions or refers to a mental operation applied to the stimulus (e.g., disassembling, re-assembling, turning, distorting, folding, etc.), or the child proposes an (imagined) act of assembling, combining, or synthesizing the stimulus with other objects or materials, to obtain a functional change of the stimulus that enables the discovery of a use. | 1. C: "If you take off these hairs (toothbrush) and then put such a brush on, and also paper on, then you can make a mouse."<br>2. C: "If you attach a lot of balloons on it (basket), which keeps floating, a lot a lot, then you can sit in there just like a hot air balloon." |
| 2b. Creating an imaginary scene | The child mentions or refers to creating mentally a scene which accompanies a use; this is often an imaginary scene wherein the use is or has been taking place; the scene might be related to the child's (memory of) pretend-play and add imagination, vividness, fantasy, or humor to the thinking process. | 1. C: "You take a spoon, and then you take a plate, and then you go to eat like a princess."<br>2. C: "You hold the wash cloth just like it is a baby, and then you do 'kukjekukje' (the child was likely petting the baby)." |
| 2c. Mentioning stimulus properties | The child explicitly mentions, refers, or points to a property or several properties of the stimulus which associatively, functionally, or conceptually relate to a use.<br>Or: The child answers affirmatively to questions of the experimenters whether the child has focused on particular properties of the stimulus objects to a use. | 1. C: "Umbrella can be used as a boat because it keeps floating."<br>2a. C: "You can (use a brick to) make that (pointing to the frame of the window)."<br>2b. T: "How do you think about that? Because they are both red?"<br>2c. C: "Yes." |
| 2d. Imagining stimulus properties | The child explicitly mentions, refers, or points to a property or several properties of the stimulus that are implicated in the mental image of the stimulus or other related objects of a use. It could be that, (1) the child has imagined a non-existing property/properties of a stimulus; (2) the child has attempted to change or transform a property/properties of a stimulus or a relevant objects of a use. And these properties are associatively, functionally, or conceptually related to a use. | 1. C: "(Use a spoon) to catch a mini dinosaur." (Here the size of dinosaur is minimized in order to fit the small size of a spoon.)<br>2. C: "If (fishnet) is very very long, then you can catch a cloud." (Here the size of the fish net is changed in order to reach the distanced cloud.)<br>3. C: "The broom has feathers (the lower part). " (Here the feathers are imagined properties of the broom in the child's mind.) |
| 2e. Mentioning elements in the test environment | The child names, refers, or points to a particular element or several elements perceived in the test environment (e.g., a picture on the wall, a book on the shelf, a pair of scissors on the desk, …) that are associatively, functionally, or conceptually related to a use.<br>Or: The child answers affirmatively to the experimenter when asked if environmental elements have influenced his or her generating of uses. | 1. C: "A brick can be used to build a chimney, because I see there is a chimney [points to a house seen through the window of the testing room]."<br>2a. C: "You can use it (fish net) to catch stars."<br>2b. T: "There are stars on your shirt, he."<br>2c. C: "Yeah."<br>2d. T: "Is it because of that, so that you think you can use it to catch stars?"<br>2e. C: "Yes." |

the number of stimuli) to be included in the child-level regression analyses to examine the associations with the AUT overall fluency and originality scores. Second, for all thinking process categories, the scores (whether a particular process was present or not) at the use level, nested within stimuli and children, were included in multilevel regression analyses to examine the serial order effects.

#### *2.2.3. Inter-coder reliability*

To establish inter-coder reliability, 10% of the transcripts (8 children) were randomly selected and double-coded by the first and the second author (see Appendix B for an overview of the inter-coder agreement). Five process categories showed moderate to excellent agreement ($\kappa = .56$–$.91$) and were used for further coding, whereas the process category Mentioning stimulus properties was excluded due to poor inter-coder agreement ($\kappa = .33$).

### *2.3. Procedures*

The AUT task was administered individually in a separate room in the participating schools during school hours. Children were picked up from their classrooms during breaks or seatwork to be tested. Six students of the master's program Clinical Child, Family and Education Studies at Utrecht University, with extensive previous training in child assessments as part of their course requirements, were trained as experimenters to conduct the tests. They carefully read the instruction manual of the test, watched example test sessions, and read corresponding transcripts. Afterwards, the first author supervised a number of test sessions in the field conducted by each experimenter and gave additional feedback where needed, until the quality of test administration was sufficient (two to three sessions for most experimenters). Throughout the field work period, the first author was present at the test sessions of all experimenters on a number of occasions to ensure the tests were administered correctly and in a consistent fashion across experimenters.

### *2.4. Data analyses*

#### *2.4.1. Missing data and final sample*

Data of twenty-three children were missing or had to be excluded from further analysis. Four children did not attend the test session, and for three children the video recordings of the test session were lost due to technical problems. All transcripts were critically reviewed for test administration quality. Through this procedure, we found that one experimenter had insufficiently prompted children to explain their thinking processes, which was a core element of the test. Therefore, data of 16 children who were tested by this experimenter were excluded. Finally, data of 84 children (40 boys and 44 girls; Age: $M = 4.44$ years, $SD = 0.26$, range = 4.02–5.10) were available for further analyses.

#### *2.4.2. Analysis plan*

In order to obtain a general impression of DT abilities in children at four years of age (Hypotheses 1 and 2), descriptive statistics of both the DT ability scores as determined with the AUT and the frequency of occurrence of the thinking process categories were computed. Subsequently, to investigate how different DT processes relate to originality and fluency (Hypothesis 3), we applied hierarchical multiple regression analyses on the data at the child level, following the approach of the study of Gilhooly et al. (2007). As a first step, we regressed children's originality and fluency scores on the frequencies of occurrence of the DT processes. As a second step, in both the fluency and the originality regression models, the frequencies with which children were prompted to generate more ideas (more-ideas prompts) and to explain their thinking processes (thinking-process prompts) were included as covariates, considering that these prompts might have had a direct influence on children's responses to the test. As a third step, for originality only, we ran an additional model in which fluency was also included as a covariate in order to control for the possible confounding effect of this variable, as higher fluency increases the chance of coming up with original ideas (for a discussion on the relation between fluency and originality, see Dumas & Dunbar, 2014). The multiple regression analyses were conducted in SPSS 24.0.

Next, multilevel regression analysis was conducted to test the serial order effects in the originality scores of generated uses (Hypothesis 4) and DT processes (Hypothesis 5). For testing Hypothesis 4, the *originality of generated uses* (proportional data at the level of generated uses) was the outcome variable. Predictors were the *rank number of a particular use in the sequence of all generated uses for a particular stimulus* (ordinal data at the level of generated uses; for the sake of convenience, we refer to this variable as *use rank*) and DT process scores (dichotomous data at the use level: coded 1 [*present*] or 0 [*absent*] for each DT process separately). More specifically, because the originality scores of the generated uses were proportions, we used multilevel logistic regressions with a binomial distribution. The following models were run: (1) M1, the intercept-only model; (2) M2, a model with linear and quadratic effects of use rank; (3) M3, a model with the fixed effects of DT processes; and (4) a model with the linear and quadratic effects of use rank but now with only the DT processes that were significant predictors of originality in the previous model (i.e., M3) to obtain a parsimonious model.

To test Hypothesis 5 on the serial order effect of DT processes, multilevel logistic regression models were applied again but now with a Bernoulli distribution, as the DT processes as outcome variables were dichotomous at the level of generated uses. For each thinking process, an intercept-only model (M5) and a model with the linear and quadratic effects of use rank (M6) were fitted. All multilevel models were fitted in SuperMix version 2.1 (Hedeker, Gibbons, du Toit, & Cheng, 2008) with adaptive quadrature estimation (20 quadrature points). Predictor variables were added to the models uncentered.

## 3. Results

### 3.1. Descriptive statistics and correlations between child-level measures

Tables 2 and 3 present the descriptive statistics and correlations of fluency, originality, and the DT process measures aggregated to the child level. Table 4 presents the number of uses that were explained by children in terms of specific thinking processes and how these references were distributed in the sequence of generated uses. As shown in the Table 4, the DT process Retrieval or recall of prior knowledge or experience was most frequently mentioned and also occurred early in the thinking flow, whereas the other DT processes were mentioned far less frequently and occurred much later.

### 3.2. Multiple regression analyses

Table 5 presents the results of the multiple regression analyses with fluency and originality as dependent variables (child-level data). For fluency, all five DT processes were significant predictors in a model without covariates (Model 1). When controlling for the frequencies of thinking-process and more-idea prompts as provided by the experimenter (Model 2), the DT processes Retrieval or recall of prior knowledge or experience and Imagining stimulus properties were no longer predictive of fluency, and the effect sizes of the other DT processes were much smaller. Note that especially the frequency of more-idea prompts was a relatively strong significant predictor of fluency. For originality, all DT processes, except Retrieval or recall of prior knowledge or experience, were significant predictors. After controlling for the frequencies of prompts, Creating an imaginary scene was no longer a significant predictor. When additionally controlling for fluency (Model 3), only the process Performing mental operations on the stimulus remained predictive of originality. Fluency was the strongest significant predictor of originality in this model.

### 3.3. Multilevel logistic regressions

Table 6 presents the results of multilevel logistic regressions with the originality of uses as the dependent variable.

#### 3.3.1. Serial order effect in originality

Compared to the intercept-only model (M1), model fit significantly improved when the linear and quadratic effects of use rank were included (M2), as is reflected in the change in deviance between the models. The linear and quadratic effects of use rank on originality were significant. Thus, the expected serial order effect in the originality of generated uses was confirmed. The positive sign of the linear effect and the negative sign of the quadratic effect indicate increasing originality with increasing rank number in the sequence of generated uses, but with the increase in originality leveling off and even decreasing again towards the end of the process of idea generation (see Fig. 2).

#### 3.3.2. Relating DT processes to originality

Adding all five DT processes as predictors to the model did not significantly improve the model fit (M3 versus M2) and only the DT process Performing mental operations on the stimulus was a significant predictor. Trimming the model by including only the process of Performing mental operations on the stimulus as predictor resulted in improved model fit (M4 versus M2), and this DT process remained a significant predictor.

#### 3.3.3. Serial order effects in the occurrence of DT processes

Table 7 presents the results of the multilevel logistic regressions with the DT processes as dependent variables.

Adding the linear and quadratic effects of use rank significantly improved model fit compared to the intercept-only models for all DT processes (M6 versus M5), except for Retrieval or recall of prior knowledge or experience. For the other DT processes, there was a significant positive linear and a negative quadratic effect of use rank (not significant for Mentioning elements in the test environment),

**Table 2**
Descriptive Statistics of Child-level Measures of Divergent Thinking (DT) and Prompts ($N = 84$).

| Variables | *M* | *SD* | Skew. | Kurt. | Min. | Max. |
|---|---|---|---|---|---|---|
| **DT ability measures (mean scores per stimulus):** | | | | | | |
| Fluency | 2.67 | 1.26 | 1.55 | 4.73 | 0.50 | 8.50 |
| Originality | 0.91 | 0.47 | 1.09 | 1.83 | 0.21 | 2.67 |
| **Frequency of DT processes (mean counts per stimulus):** | | | | | | |
| Retrieval or recall of prior knowledge or experience | 0.97 | 0.60 | 0.51 | −0.37 | 0 | 2.50 |
| Performing mental operations on the stimulus | 0.078 | 0.19 | 4.02 | 21.76 | 0 | 1.33 |
| Creating an imaginary scene | 0.052 | 0.16 | 3.72 | 13.97 | 0 | 0.83 |
| Imagining stimulus properties | 0.048 | 0.14 | 4.29 | 20.66 | 0 | 0.83 |
| Mentioning elements in the test environment | 0.048 | 0.11 | 2.36 | 5.14 | 0 | 0.50 |
| **Frequency of prompts provided by the experimenters (mean counts per stimulus):** | | | | | | |
| Thinking-process prompts | 1.32 | 0.54 | 0.13 | −0.39 | 0.00 | 2.67 |
| More-idea prompts | 2.36 | 0.82 | 0.52 | 0.29 | 0.33 | 4.50 |

**Table 3**
Correlations between Child-Level Measures of Divergent Thinking (DT) and Prompts ($N$ = 84).

| Variables | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| 1. Fluency | .80** | .31** | .45** | .44** | .29** | .51** | .41** | .64** |
| 2. Originality | | .24* | .57** | .37** | .37** | .50** | .42** | .57** |
| 3. Retrieval or recall of prior knowledge or experience | | | .16 | .22* | −.13 | .07 | .66** | .15 |
| 4. Performing mental operations on the stimulus | | | | .14 | .24* | .30** | .29** | .32** |
| 5. Creating an imaginary scene | | | | | .12 | .26* | .30** | .35** |
| 6. Imagining stimulus properties | | | | | | .13 | .11 | .26* |
| 7. Mentioning elements in the test environment | | | | | | | .25* | .50** |
| 8. Thinking-process prompts | | | | | | | | .30** |
| 9. More-idea prompts | | | | | | | | |

*$p < 0.05$. **$p < 0.01$.

similar to the findings for originality. Note that, although the linear and quadratic effects of use rank did not significantly predict the logit probability of the presence of the process Retrieval or recall of prior knowledge or experience, Fig. 3 shows an initial high level of references to this process, but a decrease at later stages.

## 4. Discussion

The present study aimed to increase our understanding of the process of divergent thinking (DT) in young children. More specifically, we investigated four-year-olds' explanations of their thinking processes during a widely used task to measure divergent thinking abilities, the Alternative Uses Task (AUT). For this purpose, the AUT was administered in dialogical sessions in which trained experimenters prompted children to come up with more ideas and encouraged them to explain how they arrived at these ideas. Children's utterances were transcribed and a coding scheme was developed based on previous work with adults (Gilhooly et al., 2007) to code the verbalizations, focusing specifically on children's utterances that directly or indirectly referred to the thinking processes underlying the generation of alternative uses in the AUT. To the best of our knowledge, this is the first study that addressed the process of DT in children at such a young age.

Our main findings are as follows: (1) When children explained how they arrived at a particular use, by far most explanations were indicative of associative processes involving semantic or episodic long-term memory, which are thought to be largely automatic, bottom-up, and relatively effortless (Beaty & Silvia, 2012; Beaty et al., 2014; Benedek et al., 2012; Gilhooly et al., 2007; Hass, 2017a, 2017b; Mednick, 1962). (2) Children's utterances were occasionally also indicative of other thinking processes, which are considered more effortful and executive in nature (e.g., Beaty, Benedek, Silvia, & Schacter, 2016; Lee & Therriault, 2013; Radel et al., 2015; Silvia, 2015; Gilhooly et al., 2007). These include *Performing mental operations on a stimulus* (e.g., disassembling, re-assembling, turning, distorting, folding), *Creating an elaborate scene* in which the stimulus is used in an imagined realistic context, *Mentioning stimulus properties* (e.g., focusing on particular properties of the stimulus), *Imagining stimulus properties* (e.g., imagining non-existent properties or imagining to change existent properties of the stimulus), and *Mentioning elements in the test environment* (switching attention from the stimulus to the immediate environment for additional cues to generate a new use of the stimulus). (3) Both associative and executive thinking processes were predictive of the number of generated uses after controlling for prompting, that is, *fluency*, as one dimension of DT ability measured with the AUT. (4) The only thinking process that was uniquely predictive of *originality*, as another dimension of DT ability, after controlling for both fluency and prompting, was Performing mental operations on the stimulus. Finally, (5) the four-year-olds in our study already showed a serial order effect in the originality of their responses, indicating that uses that were generated later in the sequence of the responses to a particular stimulus were generally more original. Similar serial order effects were found for the executive thinking processes as well: the later in the sequence a particular use was generated, the more likely it was that children would explain how they came up with this use by referring to these processes. Although children's references to associative memory-based processes showed an overall high initial probability and a decline after the first few mentioned uses, this reversed serial order effect was not significant, probably due to the high overall frequency of children's references to memory-based processes.

### *4.1. DT processes and fluency*

All DT processes were shown to be predictive of fluency in the multiple regression analyses without covariates. After controlling for the frequency of prompts provided by the experimenter, the DT processes Retrieval or recall of prior knowledge or experience and Imagining stimulus properties were no longer significant predictors of fluency. Conversely, the frequency of prompts to encourage children to come up with more ideas appeared to be a significant and strong predictor of fluency. Note that the prompts probably affected the generating of ideas *through* stimulating children's thinking process, which explains the shared variance with the DT processes. The strong predictive effect of prompts and the resulting smaller (mostly non-significant) effects for the DT processes after controlling for the prompts, therefore, may reflect that DT processes mediate the association between prompts and fluency. Furthermore, in contrast to the finding of Gilhooly et al. (2007) among adults that only associative processes involving long-term memory predicted fluency, the present study found that several executive DT processes (Performing mental operations on the stimulus, Creating an imaginary scene, and Mentioning elements in the test environment) were significant predictors of fluency, even after controlling for prompting. A possible explanation for this discrepancy is that we worked with interactive dialogues, while in the study

**Table 4**
Number of Uses and Counts of Divergent Thinking (DT) Processes by Use Rank.

| Use rank | *N* (use) | Retrieval or recall of prior knowledge or experience | Performing mental operations on the stimulus | Creating an imaginary scene | Imagining stimulus properties | Mentioning elements in the test environment | *N*(use not explained) |
|---|---|---|---|---|---|---|---|
| 1 | 482 | 191 | 1 | 0 | 1 | 3 | 286 |
| 2 | 344 | 129 | 10 | 4 | 7 | 2 | 204 |
| 3 | 214 | 70 | 13 | 4 | 7 | 4 | 123 |
| 4 | 122 | 45 | 5 | 6 | 3 | 5 | 64 |
| 5 | 66 | 28 | 3 | 7 | 1 | 5 | 26 |
| 6 | 39 | 12 | 2 | 2 | 2 | 1 | 21 |
| 7 | 25 | 7 | 2 | 0 | 2 | 3 | 13 |
| 8 | 18 | 2 | 1 | 1 | 1 | 0 | 13 |
| 9 | 10 | 2 | 1 | 1 | 0 | 0 | 7 |
| 10 | 5 | 1 | 0 | 0 | 0 | 0 | 4 |
| 11 | 6 | 0 | 1 | 1 | 0 | 0 | 4 |
| 12 | 2 | 0 | 0 | 0 | 0 | 0 | 2 |
| 13 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| *N*(use) | 1334 | 487 | 39 | 26 | 24 | 24 | 767 |
| *M*(use rank) | 2.57 | 2.35 | 3.82 | 4.58 | 3.63 | 4.33 | 2.53 |
| *N*(child) | 84 | 80 | 19 | 11 | 14 | 17 | 84 |

*Note.* The sum of the numbers of uses coded to all DT processes and the number of unexplained uses is larger than the total number of generated uses (i.e., 1334). This is due to the fact that sometimes more than one DT process was coded in relation to one use (this occurred for 32 uses). Note that 34 children reported other DT processes than the associative process Retrieval or recall of prior knowledge or experience.

**Table 5**
Hierarchical Multiple Regression Analyses with Fluency and Originality as Outcome Variables and Divergent Thinking Processes as Predictors ($N$ = 84).

| Predictor | Fluency | | | | Originality | | | |
|---|---|---|---|---|---|---|---|---|
| | $\Delta R^2$ | *B* | *SE B* | β | $\Delta R^2$ | *B* | *SE B* | β |
| **Model 1** | .51*** | | | | .56*** | | | |
| (Constant) | | 1.73 | 0.20 | | | 0.59 | 0.07 | |
| Retrieval or recall of prior knowledge or experience | | 0.46 | 0.18 | .22** | | 0.12 | 0.06 | .16+ |
| Performing mental operations on the stimulus | | 1.53 | 0.57 | .23** | | 0.91 | 0.20 | .38*** |
| Creating an imaginary scene | | 1.97 | 0.65 | .26** | | 0.52 | 0.23 | .18* |
| Imagining stimulus properties | | 1.68 | 0.73 | .19* | | 0.77 | 0.26 | .24** |
| Mentioning elements in the test environment | | 3.86 | 1.00 | .33*** | | 1.29 | 0.35 | .30*** |
| **Model 2** | .08** | | | | .04* | | | |
| (Constant) | | 0.63 | 0.37 | | | 0.28 | 0.13 | |
| Retrieval or recall of prior knowledge or experience | | 0.39 | 0.21 | .18+ | | 0.06 | 0.08 | .08 |
| Performing mental operations on the stimulus | | 1.24 | 0.53 | .19* | | 0.82 | 0.20 | .34*** |
| Creating an imaginary scene | | 1.42 | 0.62 | .19* | | 0.37 | 0.23 | .13 |
| Imagining stimulus properties | | 1.16 | 0.70 | .13 | | 0.61 | 0.26 | .19* |
| Mentioning elements in the test environment | | 2.23 | 1.02 | .19* | | 0.85 | 0.37 | .20* |
| Thinking-process prompts | | 0.03 | 0.25 | .01 | | 0.08 | 0.09 | .09 |
| More-idea prompts | | 0.54 | 0.14 | .35*** | | 0.13 | 0.05 | .23* |
| **Model 3** | | | | | .13*** | | | |
| (Constant) | | | | | | 0.15 | 0.11 | |
| Retrieval or recall of prior knowledge or experience | | | | | | −0.02 | 0.07 | −.02 |
| Performing mental operations on the stimulus | | | | | | 0.57 | 0.17 | .23** |
| Creating an imaginary scene | | | | | | 0.07 | 0.20 | .03 |
| Imagining stimulus properties | | | | | | 0.37 | 0.22 | .12+ |
| Mentioning elements in the test environment | | | | | | 0.39 | 0.32 | .09 |
| Thinking-process prompts | | | | | | 0.07 | 0.08 | .08 |
| More-idea prompts | | | | | | 0.02 | 0.05 | .03 |
| Fluency | | | | | | 0.21 | 0.04 | .56*** |

$^{+}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

**Table 6**
Multilevel Logistic Regression Analyses with Originality of Use as Outcome Variable, and the Linear and Quadratic Effects of Use Rank and the Occurrence of Divergent Thinking (DT) Processes as Predictors ($N$ = 84, with 487 units at the object level and 1334 units at the use level).

| Models | M1: Intercept-only | | M2: Effects of use rank | | M3: Effects of DT processes | | M4: Parsimonious M3 | |
|---|---|---|---|---|---|---|---|---|
| Fixed part | Coeff. (s.e.) | Odds Ratio | Coeff. (s.e.) | Odds Ratio | Coeff. (s.e.) | Odds Ratio | Coeff. (s.e.) | Odds Ratio |
| Intercept | −5.17 (0.04)*** | 0.006 | −5.68 (0.11)*** | 0.003 | −5.69 (0.12)*** | 0.003 | −5.67 (0.11)*** | 0.003 |
| **Use-level predictors** | | | | | | | | |
| Linear effect: Use rank | | | 0.26 (0.06)*** | 1.30 | 0.23 (0.06)*** | 1.26 | 0.25 (0.06)*** | 1.28 |
| Quadratic effect: Use rank | | | −0.018 (0.006)** | 0.98 | −0.016 (0.006)* | 0.98 | −0.017 (0.006)** | 0.98 |
| Retrieval or recall of prior knowledge or experience | | | | | 0.079 (0.084) | 1.08 | | |
| Performing mental operations on the stimulus | | | | | 0.54 (0.18)** | 1.72 | 0.52 (0.18)** | 1.68 |
| Creating an imaginary scene | | | | | 0.26 (0.24) | 1.30 | | |
| Imagining stimulus properties | | | | | 0.17 (0.26) | 1.18 | | |
| Mentioning elements in the test environment | | | | | 0.20 (0.26) | 1.23 | | |
| **Random part** | | | | **Variance (s.e.)** | | | | |
| Scale (variance dispersion)[a] | 0.54 | | 0.51 | | 0.51 | | 0.51 | |
| $\sigma^2_{object\ level}$ | 0.0002 (0.0008) | | 0.0002 (0.0007) | | 0.0002 (0.0007) | | 0.0002 (0.0007) | |
| $\sigma^2_{childlevel}$ | 0.00 (0.0004) | | 0.00 (0.0003) | | 0.00 (0.0003) | | 0.00 (0.0003) | |
| **Deviance of the model** (*df*) | 7745.99 (3) | | 7711.53 (5) | | 7701.26 (10) | | 7704.06 (6) | |
| **Change in deviance** (vs. a previous model, Δ *df*) | | | 34.46*** (vs. M1, 2) | | 10.27[a] (vs. M2, 5) | | 7.47*** (vs. M2, 1) | |

*Note.* M1 ~ M4 = Model 1 ~ Model 4. Deviance = −2 × log-likelihood. Extended models based on M3 and M4 including the frequencies of prompts given by the experimenter were fitted. However, the results showed that neither the model fits were improved nor the predicting patterns of predictors were changed.
$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.
[a] The estimations of 'scale' in the current study are smaller than 1, indicating that the observed variance of the outcome measure originality is smaller than the theoretical variance (i.e., $\pi^2/3$) of a binomial distribution.

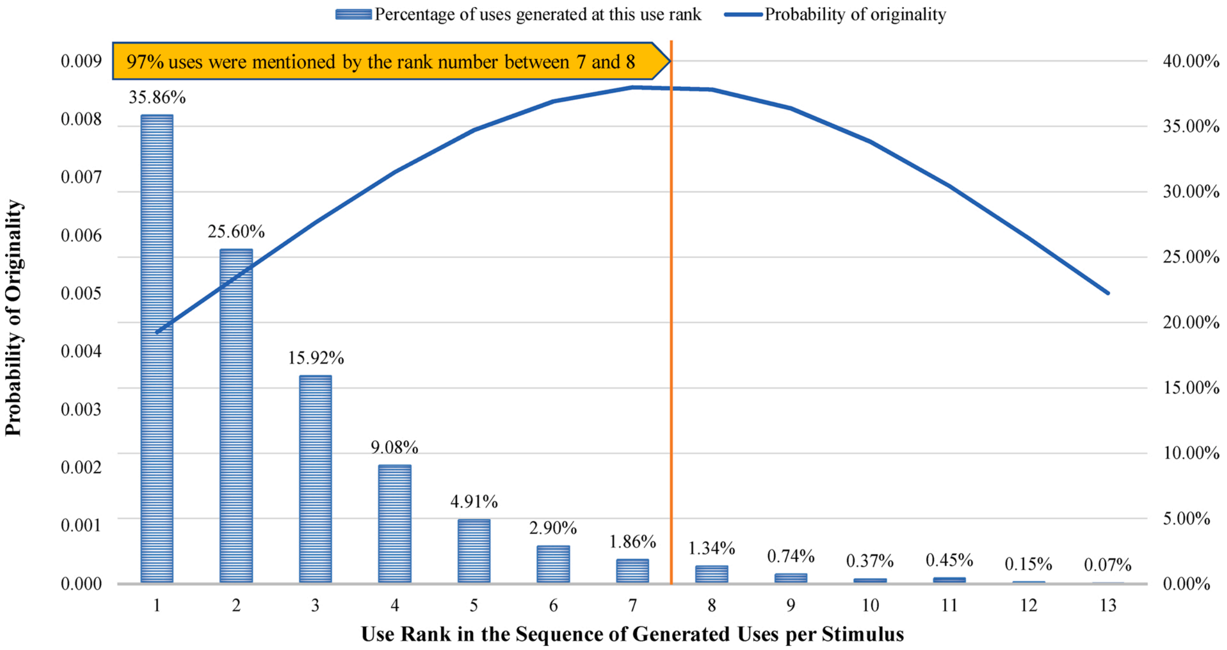


**Fig. 2.** The serial order effect of use originality: The probability of use originality by use rank in the sequence of generated uses (based on results of model M4 in Table 6 while the value of predictor Performing mental operations on the stimulus was set to 0). The columns show the percentage of uses generated at each use rank in the sequence of generated uses.

of Gilhooly et al. (2007) participants were requested to generate ideas spontaneously. Possibly, children were encouraged to spend extra effort on the task which they would not have done spontaneously. After the exhaustion of the 'easier' automatic associative process, this may have triggered the executive DT processes to supported children generating more ideas.

### 4.2. DT processes and originality

The analyses with originality as dependent variable revealed that, when controlling for prompting and fluency, only the DT process Performing mental operations on the stimulus remained uniquely predictive of originality whilst fluency was the only predictive control variable. These results may reflect that the effect of prompting on originality is also mediated by DT processes (as discussed above for fluency) or by DT processes via fluency, with higher fluency increasing the chance that original uses are mentioned (Bijvoet-van den Berg & Hoicka, 2014; Willse et al., 2008). The large shared variance of DT processes with prompting and fluency explains why most DT processes as such did not significantly predict additional variance in the model with these covariates. The current results contrast with the findings of Gilhooly et al. (2007), who found that also memory processes predicted originality. There are two possible explanations. Memory processes may increase fluency and thereby increase the probability of original responses, as argued above, if fluency is not controlled as was the case in the Gilhooly et al. study. In addition, memory-processes in adults compared to children may draw upon a larger base of semantic knowledge and personal experiences, thus increasing the probability of retrieving novel ideas. As indirect support for the latter explanation, also in our models without control for fluency, the DT process Retrieval or recall of prior knowledge or experience was not a significant predictor of originality.

Performing mental operations on the stimulus remained a significant unique predictor of originality in our study after controlling for prompting and fluency. Performing mental operations on the stimulus was coded when children's utterances reflected that they mentally disassembled the stimulus, assembled additional materials to the stimulus, or applied another kind of mental operation such as rotating or folding, which led to a clear functional change of the stimulus. Similar findings have been reported in two previous studies on adults. First, Gilhooly et al. (2007) showed that the frequency of applying the process of disassembling also made a unique contribution to the prediction of university students' originality scores. Second, in a study by Forthmann et al. (2019), undergraduate students reported that, when instructed to draw as many things as possible based on abstract figures (e.g., a circle, triangle, et cetera), three thinking processes were particularly helpful for generating ideas: (1) imagining a change in perspective on common objects so as to relate them to the figures; (2) focusing on parts of the objects which may have the same shape as the given figures; and (3) combining several figures and parts of objects in order to draw one new object. These processes closely resemble the process Performing mental operations on the stimulus in the current study. This converging evidence may point to an essential characteristic of executive processes in DT: the mentally singling-out or isolating of properties and parts of a structured whole everyday object and the subsequent recombination of these parts and properties into a new structured whole. This is in concert with what Lockman (2000) proposed as an embodied, perception-action perspective on how novel tool use unfolds in children. That is, children discover tool use through an iterative process of discovering particular affordances in objects (such as their texture or graspability), which resembles the process of

**Table 7**
Multilevel Logistic Regressions with the Divergent Thinking (DT) Processes as Outcome Variables and the Linear and Quadratic Effects of Use Rank as Predictors ($N = 84$, with 487 units at the object level and 1334 units at the use level).

| DT processes | | Retrieval or recall of prior knowledge or experience | | Performing mental operations on the stimulus | | Creating an imaginary scene | | Imagining stimulus properties | | Mentioning elements in the test environment | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Models | Results | M5 | M6 | M5 | M6 | M5 | M6 | M5 | M6 | M5 | M6 |
| **Fixed part** | | | | | | | | | | | |
| | Coeff. | −0.55*** | −0.70*** | −4.50*** | −6.82*** | −6.02*** | −9.32*** | −5.99*** | −9.66*** | −4.27*** | −6.00*** |
| Intercept | Coeff. s.e. | 0.12 | 0.21 | 0.49 | 0.97 | 0.87 | 1.55 | 0.85 | 1.95 | 0.43 | 0.73 |
| | Odds ratio | 0.58 | 0.50 | 0.01 | 0.001 | 0.002 | 0.0001 | 0.003 | 0.0001 | 0.01 | 0.003 |
| **Use-level predictors** | | | | | | | | | | | |
| | Coeff. | | 0.14 | | 0.96** | | 1.44*** | | 1.44** | | 0.75** |
| Linear effect: Use rank | Coeff. s.e. | | 0.11 | | 0.30 | | 0.43 | | 0.53 | | 0.30 |
| | Odds ratio | | 1.15 | | 2.61 | | 4.20 | | 4.20 | | 2.12 |
| | Coeff. | | −0.02 | | −0.06* | | −0.10* | | −0.11* | | −0.04 |
| Quadratic effect: Use rank | Coeff. s.e. | | 0.01 | | 0.03 | | 0.04 | | 0.05 | | 0.03 |
| | Odds ratio | | 0.98 | | 0.94 | | 0.91 | | 0.89 | | 0.96 |
| **Random part** | | | | | | **Variance (s.e.)** | | | | | |
| $\sigma^2_{\text{object level}}$ | | 0.099 (0.15) | 0.13 (0.16) | 0.24 (0.66) | 0.94 (1.09) | 0.43 (0.75) | 0.91 (1.33) | 2.45 (1.83) | 5.44 (3.82) | 0.057 (0.60) | 0.01 (0.13) |
| $\sigma^2_{\text{childlevel}}$ | | 0.89 (0.22)*** | 0.88 (0.22)*** | 1.71 (0.85)* | 1.71 (0.89)$^+$ | 4.13 (2.31) | 3.99 (2.38) | 1.94 (1.38) | 1.85 (1.81) | 0.43 (0.58) | 0.31 (0.57) |
| **Deviance of the model** (M5: *df* = 3; M6: *df* = 5) | | 1639.90 | 1636.29 | 329.97 | 309.46 | 226.42 | 201.37 | 217.76 | 203.94 | 239.60 | 222.75 |
| **Change in deviance** (M6 vs. M5, Δ *df* = 2) | | | 3.61 | | 20.51*** | | 25.05*** | | 13.82** | | 16.85*** |

*Note.* Deviance = –2 × log-likelihood; M5 = Model 5 (Intercept-only model); M6 = Model 6 (Linear and quadratic effects of use rank).
$^+p < 0.10$. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.

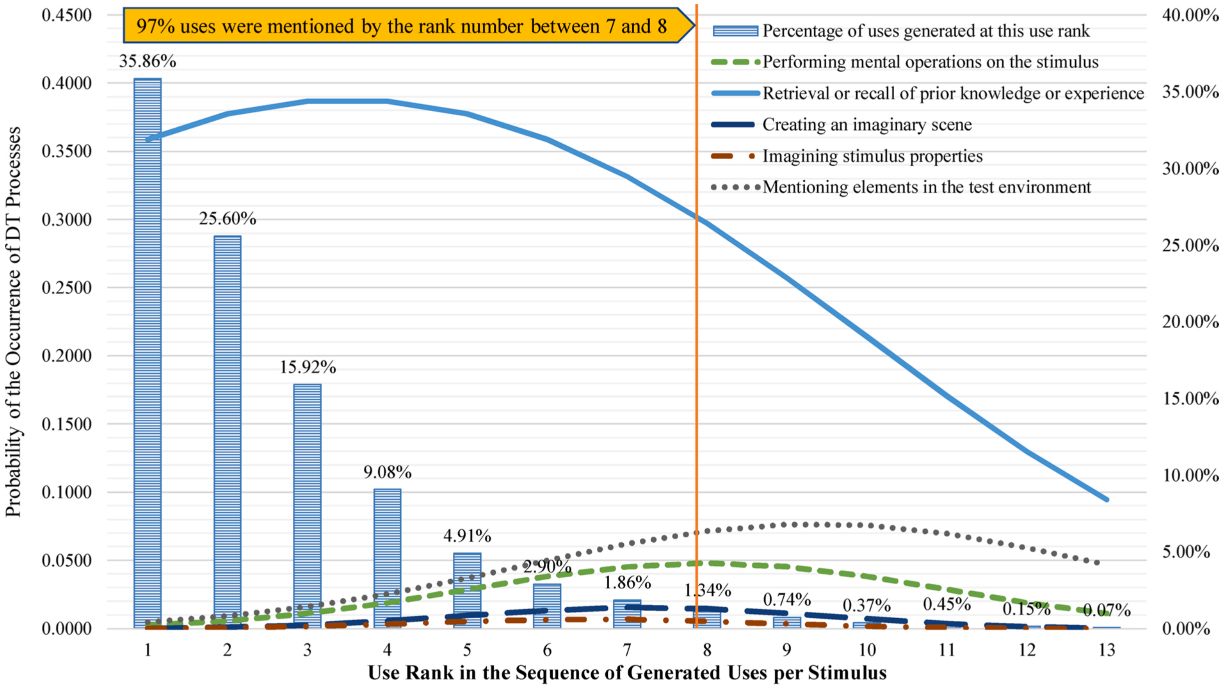


**Fig. 3.** Serial order effect of divergent thinking (DT) processes: The probability of occurrence of DT processes by use rank in the sequence of generated uses (based on results of models M6 in Table 7). Note that, the serial order effect of the process Retrieval or recall of prior knowledge or experience does not reach statistical significance. The columns show the percentage of uses generated at each use rank in the sequence of generated uses.

the singling-out of object properties, and relating these affordances to matching affordances of another object, which resembles the process of recombining properties into a new meaningful whole (see also Bourgeois, Khawar, Neal, & Lockman, 2005; Casler & Kelemen, 2005; German & Defeyter, 2000; Lockman & Kahrs, 2017).

In contrast to our expectations, none of the other executive DT processes (Creating an imaginary scene, Imagining stimulus properties, and Mentioning elements in the test environment) were found to predict originality after controlling for prompting and fluency. A possible explanation is that these processes occurred very infrequently (see Fig. 3). Future studies are needed to investigate whether these DT processes occur more frequently and are predictive of originality as children grow older.

### *4.3. Serial order effects of uses and DT processes*

In the second part of the study, we examined the serial order effects in originality and the DT processes. Based on multilevel logistic regression analysis with the rank number of a particular alternative use in the sequence of all generated alternative uses per stimulus as a quasi-time-stamped variable, a clear serial order effect was found for originality, which confirmed our expectation and is in line with previous research on adults and older children (Beaty & Silvia, 2012; Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; Christensen et al., 1957; Heinonen et al., 2016; Milgram & Rabkin, 1980; Wang et al., 2017; Ward, 1969). The results showed a curvilinear relation between the use rank and the logit probability of originality of a particular use. The logit probability of originality was found to increase from the first generated use until a maximum was reached at the use rank 7 or 8, when around 97 % of all ultimately generated uses had already been mentioned, and to decrease gradually for the remaining 3 % of generated uses thereafter. Furthermore, serial order effects were also found for all executive thinking processes. Note that the peak of the DT process Performing mental operations coincided with the peak in originality, consistent with the predictive value of this DT process for originality found in the first series of analyses. Finally, in contrast to the executive DT processes, the associative DT process Retrieval or recall of prior knowledge or experience occurred relatively frequently in relation to the first few uses. Although we found no significant serial order effect of this process, Fig. 3 suggests that the occurrence of this process actually decreased over time, coinciding with the increase of the other DT processes. Further research is needed to establish whether these two types of DT processes indeed show different serial order effects.

The current study reveals a number of key points. First, the process Retrieval or recall of prior knowledge or experience may stimulate fluency but hamper originality. A likely explanation is that the process of automatic associations, following spreading activation accounts of semantic and episodic long-term memory (Collins & Loftus, 1975; Conway & Engle, 1996), pertains to well-entrenched, strong and close connections first, thus likely generating more conventional, familiar, or thematically related associations in memory, while this process will only gradually spread to less well-entrenched, ‘weaker’, or more distal connections that are more likely to result in unconventional and truly novel alternative uses. Second, when mentioning alternative uses based on associative retrieval from long-term memory is becoming exhausted (when the strongest connections have been activated and incorporated in the subject's responses and the core semantic network is fully exploited), the more controlled, effortful thinking processes are called in (maybe in young children only if encouraged by prompts). Third, in line with our hypothesis, the gradual shift from associative to executive thinking processes, in particular the process Performing mental operations on the stimulus, co-occurs with increasing originality scores, and all processes and the originality scores reach at roughly the same time their maxima. This further suggests that originality is promoted by controlled executive processing, especially by Performing mental operations on the stimulus, in line with other research (Beaty & Silvia, 2012; Beaty et al., 2014, 2016).

### 4.4. Limitation and future research

The current study has a number of limitations. First, the use of prompts was deemed necessary to encourage children of this young age to generate ideas and to report on their thinking processes, and it helped to create a more natural dialogical situation, but it also introduced subjectivity and experimenter-variance. Although the experimenters were instructed to give prompt to all children, and always in a neutral, open way, they still had to decide when and how many prompts were given. It is to be recommended for future research to standardize the procedure of giving prompts further, as, for example, is common in learning potential tests and dynamic assessments of intellectual abilities (Resing, Elliott, & Grigorenko, 2012; Scheele et al., 2012). Also the test environment should be standardized as it was found to affect children's thinking as well in the current study. Second, asking young children to explain their thinking processes and coding their sometimes idiosyncratic utterances presents a great challenge. Only for about half of the mentioned uses, a clear and codable explanation was given, leaving unclear which DT processes were underlying the uses that were not explained by the children. Moreover, as reported in the Method section, for some DT process coding categories it appeared difficult to reach sufficient agreement between coders, and one had to be excluded from the subsequent analyses. Further refinement of the child-centred dialogical approach to DT testing, yet sufficiently standardized to reduce experimenter-variance, may support children to express their thinking more extensively and understandably. Also an in-depth content analysis of the semantic and thematic-episodic relationships between generated uses could provide more insight in the types of thinking processes children engage in when performing DT tasks. Third, in accordance with theoretical discussions in past research (e.g., Beaty et al., 2016; Beaty & Silvia, 2012; Beaty et al., 2014; Gilhooly et al., 2007), we made assumptions that the DT processes we coded involved either predominantly associative (e.g., Retrieval or recall to prior knowledge or experience) or executive (e.g., Performing mental operations on the stimulus) processes. Clearly, further studies are required to test these assumptions, for example by investigating the association between the DT process measures and measures of executive function. Finally, the current sample was drawn from kindergarten classrooms in the Dutch primary school system. It is not clear to what extent the child-centered pedagogy of the kindergarten classrooms involved in the study has influenced the current results. Given the wide variety of early childhood pedagogies across the world, caution is warranted when generalizing the findings.

### 4.5. Conclusion

The present study provides a first in-depth analysis of four-year-old children's divergent thinking processes, which shows remarkable similarities with what has been reported in studies with adults. The results confirmed the involvement of both associative and executive processes in DT, revealing, however, distinct effects on fluency and originality as outcome measures of a widely used DT task. The associative processes explained the vast majority of generated uses of the presented common objects, and typically seemed to underlie the fluency aspect of DT, whereas effortful executive processes explained a small percent of generated uses and typically seemed to underlie the originality aspect of DT. In particular, involvement of mental operations on the stimulus such as part-whole disassembling and re-assembling, perspective changing, and spatial rotation were uniquely predictive of originality. These findings not only deepen our understanding of the underlying process of idea generation, but also hold educational significance. For instance, the results of the current study provide valuable knowledge for teachers on the way young children may arrive at novel ideas. Future studies are needed to investigate whether teachers can actively encourage children to apply mental operations on available materials in order to arrive at creative solutions.

## Author statement

**Honghong Bai**: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Data curation, Writing - Original Draft, Funding acquisition, Project administration. **Hanna Mulder**: Conceptualization, Methodology, Writing - Review & Editing, Supervision. **Mirjam Moerbeek**: Formal analysis, Methodology, Validation, Writing - Review & Editing. **Evelyn Kroesbergen**: Conceptualization, Writing - Review & Editing, Funding acquisition, Supervision. **Paul Leseman**: Conceptualization, Methodology, Writing - Review & Editing, Supervision.

## Funding

This research was supported by the China Scholarship Council (grant number 201506870016).

## Acknowledgements

We are sincerely grateful to the teachers, parents, and children of the primary schools who have taken part in this study. We would also like to thank the master students who have participated in the data collection. Additionally, our great appreciation goes to fellow researchers Eveline Schoever, Marloes van Dijk, Mare Hooijdonk, and Marije Stolte for their useful comments and advices on previous drafts; and Mare Hooijdonk for her extensive support in recruiting participants.

## Appendix A. Main Types of Actions Used for Classifying Uses and Scoring Their Originality

| No. | Stimulus | | | | | |
|---|---|---|---|---|---|---|
| | Basket | Brick | Broom | Fishnet | Hand towel | Spoon |
| 1 | To put something in | To build something | To sweep | To fish | To clean somebody | For eating |
| 2 | To take something | To beat something | To clean something | To catch something | To clean something | To shovel something |
| 3 | For a picnic | To make a road | For horse riding | To put something in | To play with | For drinking |
| 4 | As a laundry basket | To play with | | To take/grab something | To dry somebody | To stir something |
| 5 | To pluck something (e.g., berry) | | | To take something out of water | | To take/grab something |
| 6 | For shopping | | | | | To take something out of a pot |
| 7 | As a pet basket | | | | | To play with |
| 8 | As a hat | | | | | |
| 9 | To play with | | | | | |

*Note.* Here we only present types of actions that were mentioned by at least four children (out of 84 children).

## Appendix B. Inter-coder Consistency Overview and Kappa's for Each Divergent Thinking Process Coding Category (137 episodes)

| Category | Frequency of coded episodes by the two coders | | | | Kappa |
|---|---|---|---|---|---|
| | yes-yes | yes-no | no-yes | no-no | |
| 1. Retrieval or recall of prior knowledge or experience | 39 | 4 | 3 | 91 | .88 |
| 2a. Performing mental operations on the stimulus | 3 | 1 | 0 | 133 | .85 |
| 2b. Creating an imaginary scene | 5 | 3 | 1 | 128 | .70 |
| 2c. Mentioning stimulus properties[a] | 1 | 0 | 4 | 132 | .33 |
| 2d. Imagining stimulus properties | 7 | 0 | 3 | 127 | .81 |
| 2e. Mentioning elements in the test environment | 2 | 3 | 0 | 132 | .56 |

[a]This category was excluded in further analyses due to unsatisfied inter-coder reliability.

## References

Ananiadou, K., & Claro, M. (2009). 21st century skills and competences for new millennium learners in OECD countries. *OECD Education Working Papers, 41*, 4–33. https://doi.org/10.1787/218525261154.

Barbot, B., Lubart, T. I., & Besançon, M. (2016). "Peaks, slumps, and bumps": Individual differences in the development of creativity in children and adolescents. *New Directions for Child and Adolescent Development, 151*, 33–45. https://doi.org/10.1002/cad.20152.

Barr, N., Pennycook, G., Stolz, J. A., & Fugelsang, J. A. (2015). Reasoned connections: A dual-process perspective on creative thought. *Thinking and Reasoning, 1*, 61–75. https://doi.org/10.1080/13546783.2014.895915.

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts, 4*, 309–319. https://doi.org/10.1037/a0029171.

Beaty, R. E., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences, 2*, 87–95. https://doi.org/10.1016/j.tics.2015.10.004.

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 7*, 1186–1197. https://doi.org/10.3758/s13421-014-0428-8.

Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 273–281. https://doi.org/10.1037/a0027059.

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence, 46*(1), 73–83. https://doi.org/10.1016/j.intell.2014.05.007.

Bijvoet-van den Berg, S., & Hoicka, E. (2014). Individual differences and age-related changes in divergent thinking in toddlers and preschoolers. *Developmental Psychology, 6*, 1629–1639. https://doi.org/10.1037/a0036131.

Bourgeois, K. S., Khawar, A. W., Neal, S. A., & Lockman, J. J. (2005). Infant manual exploration of objects, surfaces, and their interrelations. *Infancy, 3*, 233–252. https://doi.org/10.1207/s15327078in0803_3.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 2*, 77–101. https://doi.org/10.1191/1478088706qp063oa.

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science, 8*(6), 472–480. https://doi.org/10.1111/j.1467-7687.2005.00438.x.

Charles, R. E., & Runco, M. A. (2001). Developmental trends in the evaluative and divergent thinking of children. *Creativity Research Journal, 3-4*, 417–437. https://doi.org/10.1207/S15326934CRJ1334_19.

Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology, 2*, 82–88. https://doi.org/10.1037/h0045461.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 6*, 407–428. https://doi.org/10.1037/0033-295X.82.6.407.

Conway, A. R. A., & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory, 6*, 577–590. https://doi.org/10.1080/741940997.

Dumas, D., & Dunbar, K. N. (2014). Understanding fluency and originality: A latent variable perspective. *Thinking Skills and Creativity, 14*, 56–67. https://doi.org/10.1016/j.tsc.2014.09.003.

Flavell, J. H., Green, F. L., & Flavell, E. R. (2000). Development of children's awareness of their own thoughts. *Journal of Cognition and Development, 1*, 97–112. https://doi.org/10.1207/S15327647JCD0101N_10.

Forthmann, B., Wilken, A., Doebler, P., & Holling, H. (2019). Strategy induction enhances creativity in figural divergent thinking. *The Journal of Creative Behavior, 1*, 18–29. https://doi.org/10.1002/jocb.159.

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin and Review, 7*, 707–712. https://doi.org/10.3758/BF03213010.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98*, 611–625. https://doi.org/10.1348/096317907X173421.

Gralewski, J., Lebuda, I., Gajda, A., Jankowska, D. M., & Wiśniewska, E. (2017). Slumps and jumps: Another look at developmental changes in creative abilities. *Creativity. Theories – Research - Applications, 1*, 152–177. https://doi.org/10.1515/ctra-2016-0011.

Guilford, J. P. (1950). Creativity. *The American Psychologist, 9*, 444–454. https://doi.org/10.1037/h0063487.

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

Hao, N., Wu, M., Runco, M. A., & Pina, J. (2015). More mind wandering, fewer original ideas: Be not distracted during creative idea generation. *Acta Psychologica, 161*, 110–116. https://doi.org/10.1016/j.actpsy.2015.09.001.

Hass, R. W. (2016). Conceptual expansion during divergent thinking. *Proceedings of the 38th annual conference of the cognitive science society* (August).

Hass, R. W. (2017a). Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. *Memory and Cognition, 2*, 233–244. https://doi.org/10.3758/s13421-016-0659-y.

Hass, R. W. (2017b). Semantic search during divergent thinking. *Cognition, 166*, 344–357. https://doi.org/10.1016/j.cognition.2017.05.039.

Hedeker, D., Gibbons, R. D., du Toit, M., & Cheng, Y. (2008). *SuperMix: Mixed effects models*. Lincolnwood, IL: Scientific Software International.

Heinonen, J., Numminen, J., Hlushchuk, Y., Antell, H., Taatila, V., & Suomala, J. (2016). Default mode and executive networks areas: Association with the serial order in divergent thinking. *PloS One, 9*, 1–16. https://doi.org/10.1371/journal.pone.0162234.

Hoicka, E., Mowat, R., Kirkwood, J., Kerr, T., Carberry, M., & Bijvoet-van den Berg, S. (2016). One-year-olds think creatively, just like their parents. *Child Development, 4*, 1099–1105. https://doi.org/10.1111/cdev.12531.

Kirk, E., & Lewis, C. (2017). Gesture facilitates children's creative thinking. *Psychological Science, 28*, 225–232. https://doi.org/10.1177/0956797616679183.

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, 3*, 423–448. https://doi.org/10.5964/ejop.v8i3.468.

Krumm, G., Arán Filippetti, V., & Gutierrez, M. (2018). The contribution of executive functions to creativity in children: What is the role of crystallized and fluid intelligence? *Thinking Skills and Creativity, 29*, 185–195. https://doi.org/10.1016/j.tsc.2018.07.006.

Kuhn, J.-T., & Holling, H. (2009). Exploring the nature of divergent thinking: A multilevel analysis. *Thinking Skills and Creativity, 2*, 116–123. https://doi.org/10.1016/j.tsc.2009.06.004.

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence, 5*, 306–320. https://doi.org/10.1016/j.intell.2013.04.008.

Lockman, J. J. (2000). A perception-action perspective on tool use development. *Child Development, 71*, 137–144. https://doi.org/10.1111/1467-8624.00127.

Lockman, J. J., & Kahrs, B. A. (2017). New insights into the development of human tool Use. *Current Directions in Psychological Science, 26*, 330–334. https://doi.org/10.1177/0963721417692035.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 3*, 220–232. https://doi.org/10.1037/h0048850.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. London/New York: Routledge.

Milgram, R. M., & Rabkin, L. (1980). Developmental test of Mednick's associative hierarchies of original thinking. *Developmental Psychology, 2*, 157–158. https://doi.org/10.1037/0012-1649.16.2.157.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different?. Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence, 1*, 36–45. https://doi.org/10.1016/j.intell.2010.11.002.

Radel, R., Davranche, K., Fournier, M., & Dietrich, A. (2015). The role of (dis)inhibition in creativity: Decreased inhibition improves idea generation. *Cognition, 134*, 110–120. https://doi.org/10.1016/j.cognition.2014.09.001.

Resing, W. C. M., Elliott, J. G., & Grigorenko, E. L. (2012). Dynamic testing and assessment. In *Encyclopedia of the sciences of learning* (Vol. 50, pp. 1055–1058). American Library Association. https://doi.org/10.1007/978-1-4419-1428-6.

Scheele, A. F., Leseman, P. P. M., Mayo, A. Y., & Elbers, E. (2012). The relation of home language and literacy to three-year-old children's emergent academic language in narrative and instruction genres. *The Elementary School Journal, 3*, 419–444. https://doi.org/10.1086/663300.

Sharma, S., & Babu, N. (2017). Interplay between creativity, executive function and working memory in middle-aged and older adults. *Creativity Research Journal, 1*, 71–77. https://doi.org/10.1080/10400419.2017.1263512.

Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychology Review, 4*, 599–606. https://doi.org/10.1007/s10648-015-9299-1.

Silvia, P. J. (2017). Building institutions for the next 50 years of creativity research. *The Journal of Creative Behavior, 4*, 345–347. https://doi.org/10.1002/jocb.194.

Sowden, P. T., Pringle, A., & Gabora, L. (2015). The shifting sands of creative thinking: Connections to dual-process theory. *Thinking and Reasoning, 1*, 40–60. https://doi.org/10.1080/13546783.2014.885464.

Torrance, E. P. (1968). A longitudinal examination of the fourth grade slump in creativity. *The Gifted Child Quaterly, 4*, 195–199. https://doi.org/10.1016/0001-8791(92)90037-Z.

van de Kamp, M. T., Admiraal, W., & Rijlaarsdam, G. (2016). Becoming original: Effects of strategy instruction. *Instructional Science, 44*, 543–566. https://doi.org/10.1007/s11251-016-9384-y.

van Someren, M. W., Bernard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London: Academic Press. https://doi.org/10.1016/0306-4573(95)90031-4.

Wang, M., Hao, N., Ku, Y., Grabner, R. H., & Fink, A. (2017). Neural correlates of serial order effect in verbal divergent thinking. *Neuropsychologia, 99*, 92–100. https://doi.org/10.1016/j.neuropsychologia.2017.03.001.

Ward, W. C. (1969). Rate and uniqueness in children's creative responding. *Child Development, 3*, 869–878. https://doi.org/10.1111/j.1467-8624.1969.tb04548.x.

Willse, J. T., Richard, C. A., Cram, J. T., Winterstein, B. P., Barona, C. M., Martinez, J. L., … Hess, K. I. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts, 2*, 68–85. https://doi.org/10.1037/1931-3896.2.2.68.