

Divergent Thinking in Young Children

Honghong Bai

Cover design: Huang Lin (林煌) *The Dinosaurs* (恐龙)

DOI: <https://doi.org/10.33540/721>

ISBN: 978-94-6423-246-2

© 2021 Honghong Bai

The research in this dissertation was supported by the China Scholarship Council,
Grant number: 201506870016.

All rights reserved. No part of this dissertation may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronically, mechanically, by photocopy, by recordings, or otherwise, without written permission from the author.

DIVERGENT THINKING IN YOUNG CHILDREN

DIVERGENT DENKEN BIJ JONGE KINDEREN

(met een samenvatting in het Nederlands)

幼儿的发散性思维

(附中文摘要)

Proefschrift

ter verkrijging van de graad van doctor aan de

Universiteit Utrecht

op gezag van de

rector magnificus, prof.dr. H.R.B.M. Kummeling,

ingevolge het besluit van het college voor promoties

in het openbaar te verdedigen op

vrijdag 18 juni 2021 des middags te 12.15 uur

door

Honghong Bai

geboren op 11 september 1988

te Pingyang, Zhejiang, P.R. China

Promotoren:

Prof. dr. P. P. M. Leseman

Prof. dr. E. H. Kroesbergen

Copromotor:

Dr. H. Mulder

Table of Contents

Chapter 1 General Introduction	1 -
Chapter 2 Divergent Thinking in Four-Year-Old Children: An Analysis of Thinking Processes in Performing the Alternative Uses Task	9 -
Chapter 3 Serial Order Effect in Divergent Thinking in Five- to Six-Year-Olds: Individual Differences as Related to Executive Functions	35 -
Chapter 4 The Development of Divergent Thinking in Four- to Six-Year-Old Children.....	61 -
Chapter 5 The Benefits of the Learn to Think Program for Preschoolers' Creativity: An Explorative Study	81 -
Chapter 6 General Discussion.....	95 -
References	111 -
Appendices	123 -
Samenvatting (Summary in Dutch)	133 -
中文摘要 (Summary in Chinese)	139 -
Acknowledgement	145 -
About the Author	151 -
List of Publications Related to This Thesis	153 -

Chapter 1

General Introduction

“She was little, not tall enough to reach the switch of the light; she would, therefore, take the broom and grab the lower side of the stick, and then, once, twice, or sometimes with more trials, use the upper side of the stick to hit the switch and turn off the light”, a mother shares a humorous observation of how her 3-year-old daughter dealt with a daily problem. This example demonstrates perfectly that creative thinking is already present in children at a young age. As for the little girl, she has succeeded to come up with an original (at least to her) and effective idea in order to solve a problem, conforming to the two golden standards that most researchers use to define creativity (Runco & Jaeger, 2012): originality (or novelty, uniqueness) and effectiveness (or appropriateness, having value, being meaningful). In fact, if observed closely, children come up with creative ideas on a daily basis (Kaufman & Beghetto, 2009; Robson, 2014). For instance, even without having been taught so, a child might step on a chair in order to reach a book that is placed high on the shelf at school, while some other children might use a spoon as a catapult to start a “fight” with their father at the dining table. How do children come up with creative ideas? How does creative thinking develop in children? And how can we foster creativity in children? This dissertation aims to provide insights into these questions by studying a key component of creativity: divergent thinking (DT).

DT can be defined as a way of thinking that enables individuals to explore or to think in multiple directions (Guilford, 1956; Wang et al., 2017). DT is generally considered to be particularly relevant for the generation of original or creative ideas (Charles & Runco, 2001; Kuhn & Holling, 2009; Runco & Acar, 2012). In addition, DT is regarded as important for innovation and problem solving (Bijvoet-van den Berg & Hoicka, 2014), activities that often require multifaceted exploration (Jonassen, 1997). Therefore, in current society where challenges and new, complex problems continuously emerge, DT has been increasingly recognized as an important competence (Piirto, 2011; Runco & Acar, 2012) and received considerable attention in research. Tests of DT have been designed and used to measure individual differences in DT, which emerge already in children as young as 1 or 2 years of age (Hoicka et al., 2016, 2018). Further, research has also focused on describing the developmental level of DT at a certain age or over time (e.g., Daugherty, 1993; Gralewski et al., 2017; Kim, 2011; Krampen, 2012), and revealed patterns of covariation of DT with other constructs such as fluid and verbal intelligence (Silvia, 2008), executive functions (Radel et al., 2015), and real-world creative achievement (Cramond et al., 2005; Runco et al., 2010). In these studies, the research focus was predominantly on the concrete or ideational products subjects created during DT tests, for instance, the alternative uses subjects generated for common objects (e.g., a brick; Guilford, 1967). In such cases, subjects’ DT is generally measured with scores on dimensions such as fluency (the number of ideas generated), flexibility (the number of conceptual categories occurring across the generated ideas), and originality (the extent to which the ideas are original; Reiter-Palmon et al., 2019). In this dissertation we will focus mainly on originality, as originality is generally considered the core

Chapter 1

feature of divergent and creative thinking (J. K. Smith & Smith, 2017; Runco & Jaeger, 2012; Simonton, 2012; Weisberg, 2015).

The product approach as adopted in DT research is increasingly considered as limited (e.g., Glăveanu, 2013, 2018; Glăveanu et al., 2020; Kupers et al., 2019; Plucker et al., 2004; Van Dijk et al., 2019), as following this approach, the nature of the thinking processes underlying DT is often overlooked and thus, remains poorly understood. In addition, how individual characteristics influence these thinking processes in the generation of original ideas during DT is currently also understudied. Such knowledge is of great importance in deepening our understanding of what DT and creativity in general are about as well as in revealing the cognitive processes that underlie the emergence of original ideas (Glăveanu, 2013, 2018; Glăveanu et al., 2020; Kupers et al., 2019; Plucker et al., 2004).

Past research has attempted to understand the process of DT through addressing the so-called serial order effect of DT, a phenomenon that, while performing DT tasks, subjects often first generate many mundane ideas quickly and then gradually shift to the generation of truly novel ideas but at a slower pace (Christensen et al., 1957; Mednick, 1962; Milgram & Rabkin, 1980; Wang et al., 2017; Ward, 1969). The increasing originality of ideas over time during DT has been explained in two ways. First, the associative theory postulates that the generation of original ideas depends on the activation of semantic elements that are only remotely related to a given cue, through continuous automatic, associative processes over time (Beaty et al., 2014; Mednick, 1962). Second, the executive account emerging from more recent studies suggests that original ideas arise with the increasing involvement of controlled, executive thinking processes in the generation of ideas for a given cue over time (Beaty et al., 2014; Gilhooly et al., 2007). Empirical support for these explanations was mainly provided by studies that showed an association between individual differences in subjects' performance on DT tasks and either associative or executive abilities (e.g., Beaty & Silvia, 2012; Benedek, Könen et al., 2012; Hass, 2017b). However, Gilhooly and colleagues (2007) tested these explanations more directly by investigating the nature of the thinking processes that underlie DT performance through analyzing subjects' think-aloud reports during DT.

Yet, these previous studies worked only with adult samples, leaving unclear what the process of DT is in children, and in particular in kindergarteners, who have not yet received any formal education. In accordance with previous studies, formal schooling may have a negative impact on DT due to its strong focus in teaching on convergent thinking (Gralewski et al., 2016; Krampen, 2012; Russ & Wallace, 2013). Therefore, this dissertation will focus on young children from 4 to 6 years of age, who are attending kindergarten education. Specifically, this dissertation reports on empirical studies that investigated the real-time micro-development of original ideas during DT and the long-term macro-development of DT in these children, with a focus on the underlying processes of idea generation.

Compared to an approach focusing solely on product measures of DT, findings from studies on DT or creativity from a process perspective may be more informative for

educational practice. As indicated by the serial order effect, original or creative ideas do not appear out of thin air; instead, they often take time to emerge and require extensive prior exploration of a related context (Glăveanu, 2013, 2018; Van Dijk et al., 2020). This also applies to real-life scenarios. For instance, in the example presented at the beginning of this chapter, the little girl may have explored the problem in several ways (e.g., try to jump up to reach the switch) before she finally found the effective solution of using a broom to switch off the light. Specifically in the school context, knowledge about the real-time micro-development of DT and creativity in general might help teachers understand children's creative behaviors and the related inner thinking processes better in daily interactions. This could further serve as a starting point for teachers to develop more adaptive, personalized teaching strategies and instructions in order to meet every child's needs (Vygotsky, 1978). In a similar fashion, if teachers understand the long-term macro-development of DT and creativity better, in particular the predominant processes or strategies that children may adopt at a certain age, they might also become better at designing age-appropriate creativity activities, for example by designing activities that specifically focus on the training of a particular strategy (see for such an attempt Hu et al., 2011).

Another more straightforward and effective way to stimulate and foster children's DT and creativity in education is to provide preschools with well-designed intervention programs (Cheung, 2010, 2018; Chronopoulou & Riga, 2012; Craft et al., 2012; Dziedziewicz et al., 2013; Garaigordobil & Berruero, 2011; Gupta, 2009; Jamali et al., 2012; Lin, 2010; Mirzaie et al., 2009; Riga & Chronopoulou, 2014; Rizi et al., 2011; Smogorzewska, 2012, 2014). The need for such programs seems to be particularly pressing in the education system in China. Compared to education programs in other countries, Chinese education has been criticized as hindering students' creativity, likely due to its cultural roots in Confucianism which discourages behaviors that do not follow social norms (Niu et al., 2007; Niu & Sternberg, 2001, 2003; Wu, 2004; Wu & Albanese, 2010). In addition, the recent increasing emphasis on learning outcomes (Zhu & Zhang, 2008) as a consequence of the intensified competition in modern China further jeopardizes the space for activities that foster DT and creativity in schools, which also applies to kindergartens. Past research has shown that children's creativity performance, mostly assessed with DT tasks, is characterized with a steady increase with age but may temporarily decline at major transitions, for instance from home to kindergarten or from kindergarten to primary school (Barbot et al., 2016; Gralowski et al., 2016; Krampen, 2012). In China, children have been reported to show a first temporary decline in DT already at the age of about 3 years when they enroll in kindergarten (Ye & Ma, 2012), whereas in other educational contexts, the first decline in DT has often been reported to occur at the age of 6 or 7 years, when children enroll in formal education (Barbot et al., 2016; Gralowski et al., 2017; Krampen, 2012). To provide relevant input to the endeavor of implementing creativity education in the Chinese early childhood education system, this dissertation introduces a new creativity intervention program, i.e., the Learn to Think program for Preschoolers (LTT-P). The results of a field study that reviewed the benefits of

Chapter 1

this program in fostering creative thinking in 4- to 6-year-olds will be reported. The LTT-P program was specifically designed to fit in with the education system in China by taking contextual factors such as the usual classroom setting, predominant pedagogy, and common curriculum contents into account (Chinese Ministry of Education, 2012).

In sum, this dissertation has the primary aim of investigating the process of DT, addressing specifically the following questions:

- How do children come up with original ideas?
 - o What is the underlying process of DT? (Question 1)
 - o What is the role of executive processes in the real-time unfolding of original ideas? (Question 2)
- How does DT develop in children across the kindergarten years? (Question 3)

The second aim of this dissertation is to present a well-designed intervention program and to examine the benefits of this program in fostering creativity in young children in China, addressing the following question:

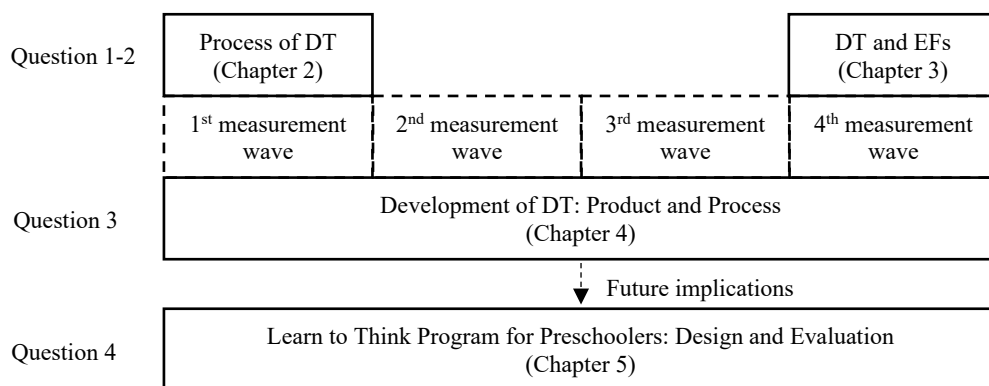
- How can we foster children's creativity? (Question 4)

The Structure of This Dissertation

The studies included in this dissertation were outcomes of two research projects. The first research project aimed to address the process of DT, answering the aforementioned Questions 1 to 3. This project adopted a longitudinal design and followed a cohort of 107 children for 1.5 years, from the age of 4 to 6 years. Children were assessed on DT with a widely applied task, the Alternative Uses Task (AUT; Guilford, 1967; Gilhooly et al., 2007), approximately every 6 months and four times in total. At the last measurement wave, children were also assessed on executive functions (EFs) using a battery of standard tests. The second project aimed to address Question 4. Within this project, a creativity intervention program was designed and assessed in a field study using a quasi-experimental design with pre- and post-tests. An overview of the structure of this dissertation, including the research design and studies presented in the chapters, is provided in Figure 1.1.

Figure 1.1

The Structure of This Dissertation



Note. DT = Divergent thinking; EFs = Executive functions.

Specifically, Chapter 2 reports on a study in which we examined the specific thinking processes involved in 4-year-olds' divergent thinking, using cohort data collected at the first measurement wave. Within this study, a novel coding scheme was developed and applied to code children's verbal reports on their thinking processes, revealing that several particular types of thinking processes were involved in children's DT. In addition, using hierarchical regression and multilevel logistic regression analysis, this study investigated how these DT processes were associated with the originality and fluency of children's generated ideas on the DT task.

Chapter 3 reports on a study in which we investigated the role of executive processes during the real-time unfolding of original ideas within a DT task in 5- to 6-year-olds, using

Chapter 1

cohort data collected at the last measurement wave. This study was unique in the aspect that domain-general executive function skills (measured by EF tests) and domain-specific (executive) DT processes were simultaneously taken into consideration to investigate how executive processes operate during DT.

In Chapter 4 a study is reported that investigated the macro-development of DT in children from the age of 4 to 6 years. In that study, the longitudinal data on children's DT was utilized. Consistent with the general emphasis of this dissertation on the process of DT, this study combined the use of common product measures (i.e., fluency and originality) and novel process measures in depicting the developmental trajectory of DT with age.

Chapter 5 reports on an intervention program, i.e., the Learn to Think program for Preschoolers (LTT-P), for fostering 4- to 6-year-old children's creativity. The program was designed for and implemented in Chinese kindergartens. Specifically, this chapter reports on the results of a quasi-experimental pretest-posttest study used to assess the potential benefits of LTT-P for the development of creativity in preschoolers.

Chapter 6 summarizes the findings of the separate studies included in this dissertation. In addition, the limitations of these studies are addressed, and the theoretical and practical implications are discussed.

Chapter 2

Divergent Thinking in Four-Year-Old Children: An Analysis of Thinking Processes in Performing the Alternative Uses Task

This chapter is published as: Bai, H., Mulder, H., Moerbeek, M., Kroesbergen, E. H., & Leseman, P. P. M. (2021). Divergent thinking in four-year-old children: An analysis of thinking processes in performing the Alternative Uses Task. *Thinking Skills and Creativity*, 40, 100814.
<https://doi.org/10.1016/j.tsc.2021.100814>

Author Contribution: *Conceptualization: H.B., E.K., P.L., and H.M.; methodology: H.B., H.M., P.L., and M.M.; investigation: H.B.; formal analysis: H.B., and M.M.; writing–original draft preparation: H.B.; writing–review and editing: E.K., P.L., H.M., and M.M.; supervision, E.K., P.L., and H.M.; funding acquisition: H.B and E.K.*

Divergent Thinking in Four-Year-Old Children: An Analysis of Thinking Processes in Performing the Alternative Uses Task

Abstract

The present study examined the divergent thinking (DT) processes of 4-year-old 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 4 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.

Keywords: divergent thinking, executive processes, isolating and recombining properties, young children, Alternative Uses Task, 4-year-old

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 (DT) as a main component of creativity, relates to the increased importance attached to creativity and DT as typical 21st century skills (Ananiadou & Claro, 2009). DT is defined as a thought process towards exploring multiple solutions (Wang et al., 2017) and is of particular importance for creativity as it enables individuals to think in multiple directions (Charles & Runco, 2001; Kuhn & Holling, 2009). Research on DT to date has shown that even 1- and 2-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 et al., 2016; Gralewski et al., 2017; Krampen, 2012; Torrance, 1968). As such, the thinking processes underlying novel idea generation on DT tasks in children are 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 4-year-olds' thinking when performing the Alternative Uses Task (AUT; Guilford, 1967)—a commonly used DT task. First, we investigated the processes of generating original ideas on the AUT in 4-year-olds. To this end, we adapted an established coding scheme that was originally designed for analyzing the process of DT in adults (Gilhooly et al., 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 et al., 1957).

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 7 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, et al., 2015; Beaty et al., 2014; Hass, 2017a; Sowden et al., 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).

Chapter 2

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 et al., 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 executive functions such as 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 et al., 2019; Krumm et al., 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, cf. Radel et al., 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, 2017a). 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.

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 4-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 et al. (1994; for an example of measuring narrative abilities in young children using prompted story [re]telling, see Scheele et al., 2012), so that children were actively prompted to give more ideas and to explain their thinking process during the AUT test. Young children, age 4 or 5 years, are not yet well capable of unsupported introspection into their mental processes (Flavell et al., 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 4 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

Chapter 2

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).

Method

Participants

The current study is part of a longitudinal research project that investigates the development of DT in children from age 4 to age 6 years. Participants were recruited from the kindergartens of four typical primary schools in the Netherlands. 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-centered 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 6 years, to prepare children for formal education. 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. Hundred seven children ranging from 3.87 to 5.10 years of age ($M = 4.44$ years, $SD = 0.26$; 49 boys and 58 girls) with informed, active parental approval were enrolled in the project. Based on available data, we estimate the positive response rate to be about 70%. 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).

Measures

The AUT task 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 prompt” from here on), e.g., “What else can you use a basket for?”, (2) to elaborate on their ideas (“explain-idea prompt”), e.g., “How do you do that?”, and (3) to explain their thinking process (“thinking-process prompt”), e.g., “How did you come up with this idea?”, “Have you done that before?”. 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 minutes (including the pretest instruction and occasional breaks). All test sessions were video-recorded.

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 Figure 2.1 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 each child for the analyses.

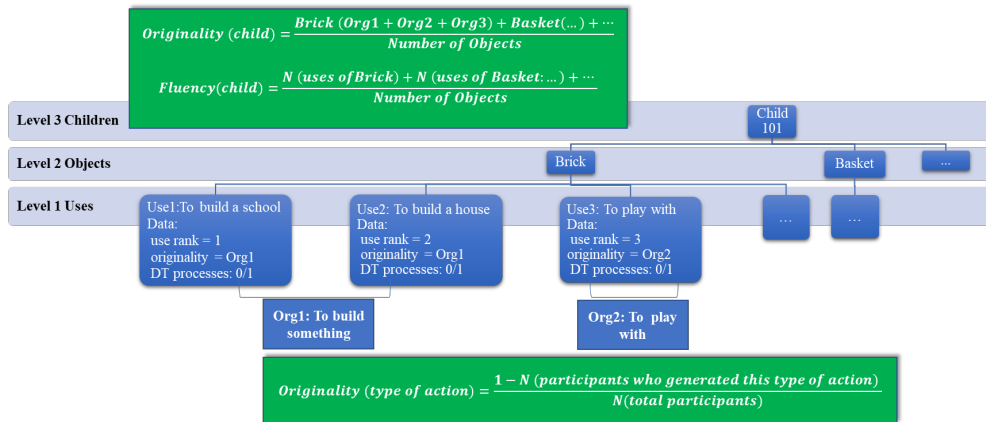
For scoring originality, the generated uses were further categorized based on the implicated type of action, in line with past research (e.g., Krampen, 2012; Van de Kamp et al., 2016). More specifically, “using a brick to build a house” and “using a brick to build a bridge” were categorized as the same *type* of action, that is, “using a brick to build something” (see Appendix 2.A for a list of action categories used for classifying 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 et al., 2015; Kirk & Lewis, 2017) and was calculated as follows: Originality of an action = $1 - (\text{The number of participants who generated this type of action} / \text{The total number of})$

Chapter 2

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 Figure 2.1). 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).

Figure 2.1

The Calculation of Fluency and Originality From the Alternative Uses Task and the Nested Data Structure of Originality and Divergent Thinking (DT) Processes



Note. 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.

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 4- to 6-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 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 2.1).

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 test 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 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.

Inter-Coder Reliability. To establish inter-coder reliability, 10% of the transcripts (eight children) were randomly selected and double-coded by the first and the second author (see Appendix 2.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 ($\kappa = .33$).

Table 2.1

Coding Scheme for Analyzing Children's Divergent Thinking (DT) Processes in the Alternative Uses Task, Adapted From Gilhooly et al. (2007)

DT Process	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. 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	1. C: "I always do that with my father." 2a. T: "Have you done that before?" 2b. C: "Yes, I have done it once."

Chapter 2

DT Process	Definition	Examples (C = Child; T = Experimenter)
	he or she did/learned it before) or if the child knew about the use from others, stories, movies, or books.	
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.	<p>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."</p> <p>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."</p>
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.	<p>1. C: "You take a spoon, and then you take a plate, and then you go to eat like a princess."</p> <p>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)."</p>
2c. Mentioning stimulus properties	<p>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.</p> <p>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.</p>	<p>1. C: "Umbrella can be used as a boat because it keeps floating."</p> <p>2a. C: "You can (use a brick to) make that (pointing to the frame of the window)."</p> <p>2b. T: "How do you think about that? Because they are both red?"</p> <p>2c. C: "Yes."</p>
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.	<p>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.)</p> <p>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.)</p> <p>3. C: "The broom has feathers (the lower part)." (Here the feathers are imagined properties of the broom in the child's mind.)</p>
2e. Mentioning elements in the test environment	<p>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.</p> <p>Or: The child answers affirmatively to the experimenter when asked if environmental elements have influenced his or her generating of uses.</p>	<p>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]."</p> <p>2a. C: "You can use it (fish net) to catch stars."</p> <p>2b. T: "There are stars on your shirt, he."</p> <p>2c. C: "Yeah."</p> <p>2d. T: "Is it because of that, so that you think you can use it to catch stars?"</p> <p>2e. C: "Yes."</p>

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.

Data Analyses

Missing Data and Final Sample

Data of 23 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.

Analysis Plan

In order to obtain a general impression of DT abilities in children at 4 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 prompt) and to explain their thinking processes (thinking-process prompt) 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

Chapter 2

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 analyses were conducted to test the serial order effects in the originality scores of generated uses (Hypothesis 4) and the occurrence of DT processes (Hypothesis 5). For testing Hypothesis 4, the *originality of generated uses* (proportional data at the use level) 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 use level; 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: (a) M1, the intercept-only model; (b) M2, a model with the linear and quadratic effects of use rank; (c) M3, a model with the fixed effects of DT processes; and (d) M4, 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 consist of dichotomous data at the use level. 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 et al., 2008) with adaptive quadrature estimation (20 quadrature points). Predictor variables were added to the models uncentered.

Results

Descriptive Statistics and Correlations Between Child-Level Measures

Tables 2.2 and 2.3 present the descriptive statistics and correlations of fluency, originality, and the DT process measures aggregated to the child level. Table 2.4 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 Table 2.4, 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.

Table 2.2*Descriptive Statistics of Child-Level Measures of Divergent Thinking (DT) and Prompts (N = 84)*

Variable	<i>M</i>	<i>SD</i>	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 (mean counts per stimulus):						
Thinking-process prompt	1.32	0.54	0.13	-0.39	0.00	2.67
More-idea prompt	2.36	0.82	0.52	0.29	0.33	4.50

Table 2.3*Correlations Between Child-Level Measures of Divergent Thinking and Prompts (N = 84)*

Variable	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 prompt								.30**
9. More-idea prompt								

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

Chapter 2

Table 2.4

Number of Uses and Counts of Divergent Thinking (DT) Processes by Use Rank

Use rank	<i>N</i> (uses)	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	<i>N</i> (uses 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
<i>N</i> (uses)	1334	487	39	26	24	24	767
<i>M</i> (use rank)	2.57	2.35	3.82	4.58	3.63	4.33	2.53
<i>N</i> (children)	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 because that sometimes more than one DT process was coded in relation to one use (this occurred for 32 uses). Thirty-four children reported other DT processes than the associative process Retrieval or recall of prior knowledge or experience.

Multiple Regression Analyses

Table 2.5 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.

Table 2.5

Hierarchical Multiple Regression Analyses With Fluency and Originality as Outcome Variables and Divergent Thinking Processes as Predictors (N = 84)

Predictor	Fluency				Originality			
	ΔR^2	B	SE B	β	ΔR^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 prompt		0.03	0.25	.01		0.08	0.09	.09
More-idea prompt		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 prompt						0.07	0.08	.08
More-idea prompt						0.02	0.05	.03
Fluency						0.21	0.04	.56***

⁺ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Multilevel Logistic Regressions

Serial Order Effect in Originality

Table 2.6 presents the results of multilevel logistic regressions with the originality of uses as the dependent variable.

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 Figure 2.2).

Relating DT Processes to Originality. Adding all five DT processes as predictors to the model did not significantly improve the model fit (M3 vs. 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 vs. M2), and this DT process remained a significant predictor.

Serial Order Effects in the Occurrence of DT Processes

Table 2.7 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 vs. M5), except for Retrieval or recall of prior knowledge or experience. For the other DT processes, there were a positive linear effect and a negative quadratic effect of use rank (not significant for Mentioning elements in the test environment), 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, Figure 2.3 shows an initial high level of references to this process, but a decrease at later stages.

Table 2.6

Multilevel Logistic Regression Analyses With Originality of the Uses as the 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)

Model	M1: Intercept-only		M2: Effects of use rank		M3: Effects of DT processes		M4: Parsimonious M3	
Fixed part	<i>Coeff.</i> (<i>s.e.</i>)	<i>OR</i>	<i>Coeff.</i> (<i>s.e.</i>)	<i>OR</i>	<i>Coeff.</i> (<i>s.e.</i>)	<i>OR</i>	<i>Coeff.</i> (<i>s.e.</i>)	<i>OR</i>
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
Linear: Use rank			0.26 (0.06)***	1.30	0.23 (0.06)***	1.26	0.25 (0.06)***	1.28
Quadratic: 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 (<i>s.e.</i>)							
Scale (variance dispersion) ^a	0.54		0.51		0.51		0.51	
$\sigma^2_{\text{object level}}$	0.0002 (0.0008)		0.0002 (0.0007)		0.0002 (0.0007)		0.0002 (0.0007)	
$\sigma^2_{\text{child level}}$	0.00 (0.0004)		0.00 (0.0003)		0.00 (0.0003)		0.00 (0.0003)	
Deviance (<i>df</i>)	7745.99 (3)		7711.53 (5)		7701.26 (10)		7704.06 (6)	
Δ deviance (Δ <i>df</i>)			34.46*** (2), vs. M1		10.27 ^a (5), vs. M2		7.47*** (1), vs. M2	

Note. Deviance = $-2 \times \log\text{-likelihood}$. *OR* = Odds ratio. Extended models based on M3 and M4 including the frequencies of prompts given by the experimenter were fitted, and the results showed that neither the model fits were improved nor the predicting patterns of predictors were changed.

^a The estimations of *scale* are smaller than 1, indicating that the observed variance of the outcome measure is smaller than the theoretical variance (i.e., $\pi^2/3$) of a binomial distribution.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 2.7

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 process	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	
	M5	M6	M5	M6	M5	M6	M5	M6	M5	M6
Fixed part										
Intercept	<i>Coeff.</i>	-0.54***	-0.70***	-4.50***	-6.82***	-6.02***	-9.32***	-5.99***	-9.66***	-4.27***
	<i>Coeff. s.e.</i>	0.12	0.21	0.49	0.97	0.87	1.55	0.85	1.95	0.43
Linear: Use rank	<i>OR</i>	0.58	0.50	0.01	0.001	0.002	0.0001	0.003	0.001	0.01
	<i>Coeff.</i>		0.14		0.96**		1.44***		1.44**	
Quadratic: Use rank	<i>Coeff. s.e.</i>		0.11		0.30		0.43		0.53	
	<i>OR</i>		1.15		2.61		4.20		4.20	
Quadratic: Use rank	<i>Coeff.</i>		-0.02		-0.06*		-0.10*		-0.11*	
	<i>Coeff. s.e.</i>		0.01		0.03		0.04		0.05	
	<i>OR</i>		0.98		0.94		0.91		0.89	
Random part										
$\sigma^2_{\text{object level}}$		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)
$\sigma^2_{\text{child level}}$		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)
Deviance ^a		1639.90	1636.29	329.97	309.46	226.42	201.37	217.76	203.94	239.60
Adeviance ^b			3.61		20.51***		25.05***		13.82**	

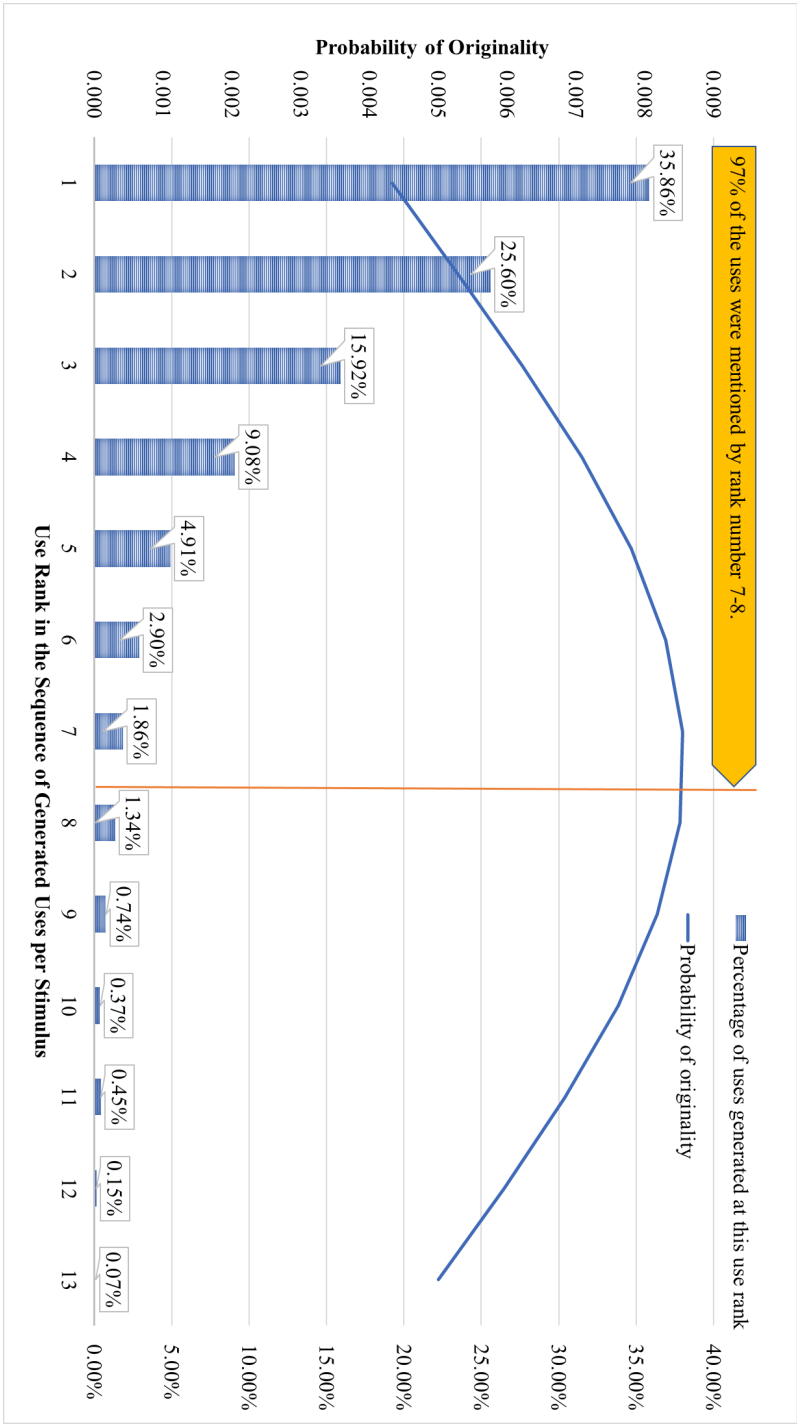
Note. Deviance = $-2 \times \log\text{-likelihood}$. *OR* = Odds ratio. M5: Intercept-only model; M6: The model with the linear and quadratic effects of use rank.

^a M5: *df* = 3; M6: *df* = 5. ^b M6 vs. M5, $\Delta df = 2$.

[†] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 2.2

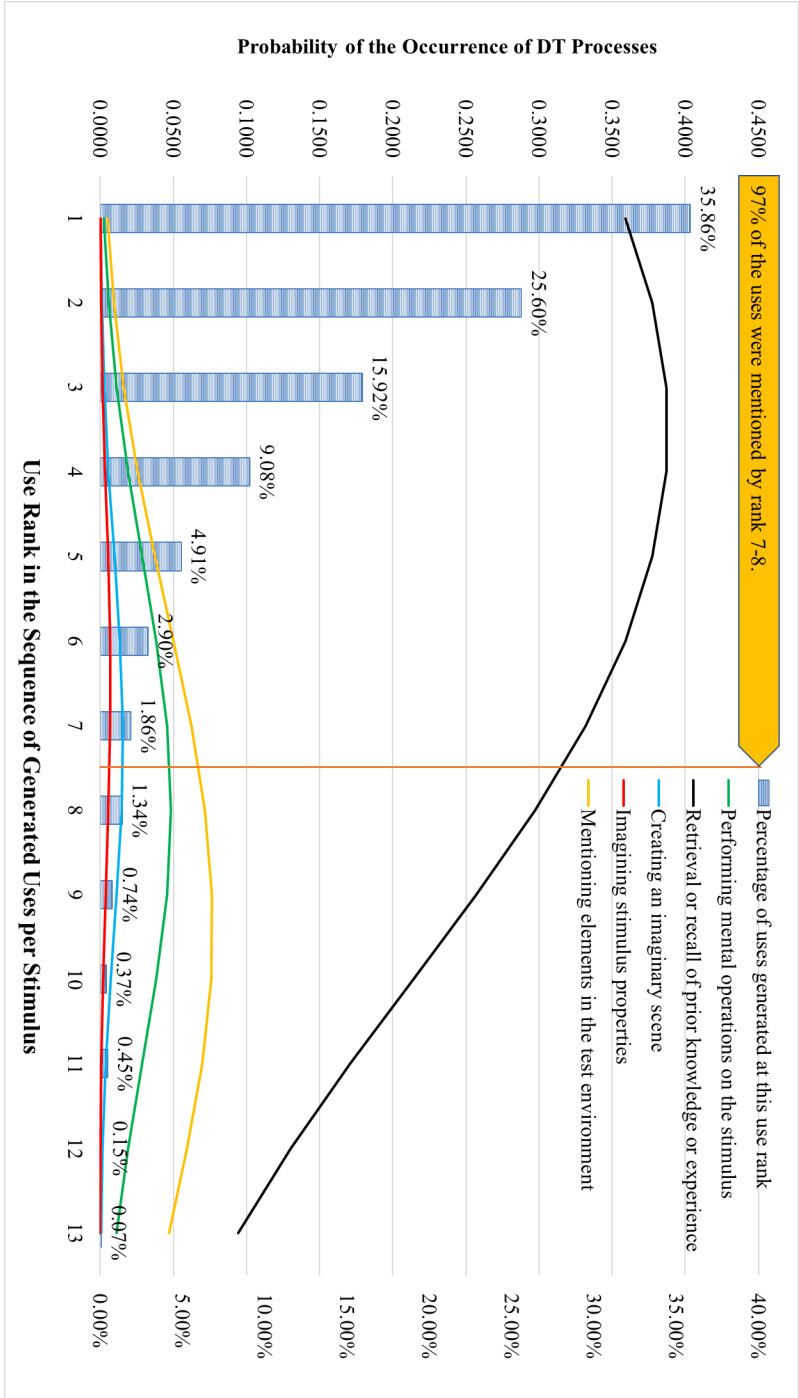
The Serial Order Effect of Originality: The Probability of Originality by Use Rank in the Sequence of Generated Uses



Note. This figure was plotted based on the results of the model M4 in Table 2.6, with the value of the predictor Performing mental operations on the stimulus set to zero. The columns show the percentage of uses generated at each use rank in the sequence of generated uses.

Figure 2.3

Serial Order Effects of Divergent Thinking (DT) Processes: The Probability of the Occurrence by Use Rank in the Sequence of Generated Uses



Note. This figure was plotted based on the results of M6 models in Table 2.7. The serial order effect of the process Retrieval or recall of prior knowledge or experience was significant. The columns show the percentage of uses generated at each use rank in the sequence of generated uses.

Discussion

The present study aimed to increase our understanding of the process of DT in young children. More specifically, we investigated 4-year-olds' explanations of their thinking processes during a widely used task to measure DT ability, the AUT task. For this purpose, the AUT was administered in dialogical sessions in which trained experimenters prompted children to come up with more ideas and 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, Könen et al., 2012; Mednick, 1962; Gilhooly et al., 2007; cf. Hass, 2017a). (2) Children's utterances were occasionally also indicative of other thinking processes, which are considered more effortful and executive in nature (e.g., Beaty et al., 2016; Gilhooly et al., 2007; Lee & Therriault, 2013; Radel et al., 2015; Silvia, 2015). These include *Performing mental operations on a stimulus* (e.g., disassembling, re-assembling, turning, distorting, folding), *Creating an imaginary 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 4-year-olds in our study already showed a serial order effect in the originality of their responses, indicating that uses 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.

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 generation 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 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.

DT Processes and Originality

The analyses with originality as the 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-based processes predicted originality. There are two possible explanations. Memory-based 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-based 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 operations 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. (2016), 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: (a) imagining a change in perspective on common objects so as to relate them to the figures; (b) focusing on parts of the objects which may have the same shape as the given figures; and (c) 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: (a) the mentally singling-out or isolating of properties and parts of a structured whole everyday object and (b) 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 grasp-ability), which resembles the process of 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 et al., 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 Table 2.4). Future studies are needed to investigate whether these DT processes occur more frequently and are predictive of originality as children grow older.

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 occurrence of DT processes. Based on multilevel logistic regression analysis with the rank number of a particular use in the sequence of all generated 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 et al., 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. As shown in Figure 2.2, the 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 the occurrence of all executive DT processes (see Figure 2.3). 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, Figure 2.3 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 executive 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 et al., 2016; Beaty & Silvia, 2012; Beaty et al., 2014).

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 prompts 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 et al., 2013; 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-centered 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, 2014, 2016; Beaty & Silvia, 2012; 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.

Conclusion

The present study provides a first in-depth analysis of 4-year-old children's DT 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

Chapter 2

deepen our understanding of the underlying processes 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.

Chapter 3

Serial Order Effect in Divergent Thinking in Five- to Six-Year-Olds: Individual Differences as Related to Executive Functions

This chapter is published as: Bai, H., Leseman, P. P. M., Moerbeek, M., Kroesbergen, E. H., & Mulder, H. (2021). Serial order effect in divergent thinking in five- to six-year-olds: Individual differences as related to executive functions. *Journal of Intelligence*, 9(2), 20. <https://doi.org/10.3390/jintelligence9020020>

Author Contribution: *Conceptualization: H.B., E.K., P.L., and H.M.; methodology: H.B., H.M., P.L., and M.M.; investigation: H.B.; formal analysis: H.B., and M.M.; writing–original draft preparation: H.B.; writing–review and editing: E.K., P.L., H.M., and M.M.; supervision, E.K., P.L., and H.M.; funding acquisition: H.B and E.K.*

Serial Order Effect in Divergent Thinking in Five- to Six-Year-Olds: Individual Differences as Related to Executive Functions

Abstract

This study examined the unfolding in real time of original ideas during divergent thinking (DT) in 5- to 6-year-olds and related individual differences in DT to executive functions (EFs). The Alternative Uses Task was administered with verbal prompts that encouraged children to report on their thinking processes while generating uses for common objects. In addition to coding the originality of each use, the domain-specific DT processes *memory retrieval* and *mental operations* were coded from children's explanations. Six EF tasks were administered and combined into composites to measure working memory, shifting, inhibition, and selective attention. The results replicated findings of a previous study with the same children but at age 4 years: (1) there was a serial order effect of the originality of uses; and (2) the process mental operations predicted the originality of uses. Next, the results revealed that both domain-general EFs and domain-specific executive processes played a role in the real-time unfolding of original ideas during DT. Particularly, the DT process mental operations was positively related to the early generation of original ideas, while selective attention was negatively related to the later generation of original ideas. These findings deepen our understanding of how controlled executive processes operate during DT.

Keywords: divergent thinking, serial order effect, executive functions, selective attention, Alternative Uses Task, 5- to 6-year-olds

Divergent thinking (DT) is a thinking process to generate original ideas by exploring many possible solutions to a particular problem (Guilford, 1956; Wang et al., 2017). As an important component of creativity, DT has been the subject of abundant research. Tests of DT have been designed and related to other constructs, revealing individual differences and patterns of covariation with, for example, fluid and verbal intelligence (Silvia 2008), executive functions (Radel et al., 2015), and real-world creative achievement (Cramond et al., 2005; Runco et al., 2010), while also showing discriminant validity (Runco & Acar, 2012). Still, DT as a process that unfolds in real time remains poorly understood, as are the relations of DT with general cognitive processes. This holds in particular with regard to the question whether DT is based on automatic, relatively effortless processes involving semantic and episodic memory (Brown, 1973; Benedek, Könen, et al., 2012; Coney & Serna, 1995; Mednick, 1962; Milgram & Rabkin, 1980), or also involves controlled, relatively effortful processes, pointing to possible involvement of executive functions (Beaty et al., 2014, 2016; Beaty & Silvia, 2012; Benedek & Fink, 2019). The current study aims to address this issue by investigating the associations between individual differences in effortful processes and the real-time unfolding of original ideas in a DT task in young children.

A line of research informative to this question has focused on the so-called serial order effect. The serial order effect refers to the phenomenon that, while performing DT tasks (e.g., generating alternative uses of a brick in the Alternative Uses Task), participants often begin with generating many ideas in a short time frame, and these early ideas are usually relatively mundane and conventional. Gradually, participants tend to switch to generating more original ideas yet at a slower pace. The serial order effect has been reported in a range of age samples, from early childhood to adulthood (Bai, Mulder, et al., 2021; Christensen et al., 1957; Mednick, 1962; Milgram & Rabkin, 1980; Wang et al., 2017; Ward, 1969).

A possible explanation for the occurrence of the serial order effect is that DT involves activating automatic associations in memory (Benedek, Könen, et al., 2012; Brown, 1973; Mednick, 1962). Following spreading activation accounts of semantic and episodic long-term memory (Collins & Loftus, 1975; Conway & Engle, 1996), this process pertains to well-entrenched near connections first, therefore likely generating more conventional, familiar, or thematically related associations in memory. Activation will then only gradually spread to less well-entrenched, “weaker”, and more distal connections that are more likely to result in unconventional and truly novel alternative uses. The latter process presumably requires sustained effort and may be facilitated by domain-general intelligence (Beaty & Silvia, 2012; Hass, 2017a). A related, but slightly different explanation for the occurrence of the serial order effect is that DT initially mainly involves generating alternative uses based on automatic associative retrieval from long-term memory, while more controlled, effortful thinking processes are called in when memory-based DT is becoming exhausted (i.e., the strongest connections have been activated and incorporated in the subject’s responses, the core semantic network is fully exploited). This results in a gradual increase in qualitatively different thinking processes no longer based on finding (remote) memory associations but on

Chapter 3

manipulating mental representations and creating new links between (parts of) representations, which may be facilitated by attention-based executive functions and working memory (Bai, Mulder, et al., 2021; Gilhooly et al. 2007). Both explanations suggest that originality in DT is promoted by controlled, effortful processing. Support for the involvement of controlled, effortful processing in DT and creativity is found in studies on adults (Beaty et al., 2014, 2016; Beaty & Silvia, 2012; Benedek & Fink, 2019).

Even though 1- and 2-year-old children are already able to think divergently (Bijvoet-van den Berg & Hoicka, 2014; Hoicka et al., 2016) and also the serial order effect has been confirmed to occur in childhood (Ward, 1969), DT as a process that unfolds in real time is not well understood in children. To begin to unravel this issue, we previously conducted a detailed analysis of domain-specific processes occurring during a DT task, the Alternative Uses Task (AUT; Guilford, 1967), in 4-year-olds (Bai, Mulder, et al., 2021). In the AUT, children are instructed to generate as many unusual uses as possible for a number of stimuli (e.g., a picture of a brick). We found the serial order effect: later uses per object were more original. In addition to the standard procedure of the AUT, we asked children to explain how they came up with a particular use. Children referred to several categories of thinking processes in their explanations and two of them stood out. First, the process of retrieving uses from memory (e.g., “you can use a toothbrush to brush your teeth because I do it every day”; similar to memory-based production in a study on adults by Gilhooly et al., 2007) supported the generation of the largest number of uses compared to other processes. Second, the process of performing mental operations on the stimuli (e.g., “you can take the hairs [of a toothbrush] off and use them as eyelashes of a doll”; similar to disassembly-use production in Gilhooly et al., 2007) was found to uniquely predict the originality of uses.

Notably, these thinking processes also showed serial order effects, although in different directions. Memory retrieval occurred frequently initially, but the occurrence decreased with the increasing rank of a use in the sequence of all mentioned uses for a particular stimulus. The process mental operations occurred very infrequently initially, but the occurrence increased with increasing rank. Visual inspection of the graphs representing the serial order effects of originality and the process mental operations revealed co-occurrence of the rise, peak, and levelling-off, suggesting that especially this process may underlie relatively original responses (Bai, Mulder, et al., 2021). We argued that, in contrast to memory retrieval based on automatically activated memory associations, the process of performing mental operations on a stimulus is executive and effortful in nature, given that performing mental operations on a stimulus involves holding an image of the object in mind while manipulating it, inhibiting the perceptual or representational salience of the object in its usual whole integrated form, and shifting attentional focus from global, holistic to local, singled-out features of the stimulus in order to operate on them.

To date, several studies have investigated the association between DT performance, such as originality, and domain-general executive function skills. As an umbrella term,

executive functions (EFs) refer to the goal-oriented regulation of one's thoughts, actions, and emotions (Moriguchi et al., 2016), involving a set of top-down, effortful mental processes. There is general agreement that there are three core EFs (Diamond, 2013; Miyake et al., 2000): (a) inhibition, which refers to the ability to suppress one's predominant behavioral tendencies, thoughts, and/or emotions, in favor of sub-dominant responses; (b) working memory, which refers to the ability and capacity to hold information in mind while mentally processing that information; and (c) shifting, which refers to the ability to adapt one's thoughts and behaviors to the changing needs of a task or situation.

Study findings regarding the association between working memory, shifting and DT are rather consistent, typically showing that with enhanced working memory and shifting ability, participants performed better on DT tasks (De Dreu et al., 2012; Kharkhurin & Wei, 2015; Lee & Theriault, 2013; Lu et al., 2017; Nusbaum & Silvia, 2011; Yeh et al., 2015; Zabelina et al., 2019; Zabelina & Robinson, 2010). Possible explanations are that a relatively large working memory capacity enables participants (a) to hold more information in mind so that complex, uncommon associations can be formed (Zabelina et al., 2019) and (b) to simultaneously hold in mind and evaluate several candidate responses to select only original ideas (Lee & Theriault, 2013). In addition, the ability to flexibly shift between the analytic, top-down and the associative, bottom-up thinking modes (Pringle & Sowden, 2017; Zabelina & Robinson, 2010) or between different task stimuli (Lu et al., 2017) may help participants to generate out-of-the-box responses.

However, findings regarding the association between inhibition and DT are mixed, with some studies showing a positive association (Benedek, Franz, et al. 2012; Benedek et al., 2014; Edl et al., 2014; Krumm et al., 2018; Rominger et al., 2017; Zabelina & Ganis, 2018) and other studies showing a negative association (Radel et al., 2015; Scibinetti et al., 2011) or no association (Stolte et al., 2019; Zabelina et al., 2019). Some scholars have argued that enhanced inhibition ability is positively related to DT performance, because inhibition is beneficial for suppressing common and already-generated ideas while staying focused on the task and ignoring distraction (Beaty & Silvia, 2012; Beaty et al. 2014; Benedek, Franz, et al. 2012; Benedek et al., 2014;). In contrast, others have proposed that reduced inhibition is conducive to DT by allowing information that is only peripherally related to the stimulus to enter working memory (Radel et al., 2015), thus broadening the scope of attention and enhancing the breadth of input for further processing (Stolte et al., 2019; Zabelina, 2018).

Only a few studies to date, involving adult samples, have investigated the relationships between domain-general EFs and the serial order effect of originality in DT, showing inconsistent findings. In one study, Wang et al. (2017) investigated how inhibition, working memory, and shifting moderated the serial order effect of the originality of uses in the AUT. Shifting, but not inhibition or working memory, was found to significantly moderate the serial order effect. Adults with higher shifting ability performed better throughout the task and showed the serial order effect (i.e., originality of uses was higher later on in the series of

mentioned uses), whereas adults with lower shifting ability generated more mundane ideas overall and did not show the serial order effect. Further, one study reported a reduced serial order effect in individuals with high EF (Beatty & Silvia, 2012), while another study showed no moderation of the serial order effect by EF (Hass, 2017). To summarize, there is some evidence that EF and DT are related, but results to date are inconsistent, and it is currently unclear how EFs affect the process of idea generation during DT tasks.

The Present Study

The present study is partly a replication of our previous study using the same sample of children but now at the age of 5 to 6 years, about 1.5 years later than in our previous study (Bai, Mulder, et al., 2021). Furthermore, the present study extends our previous work by examining the role of EFs in DT through relating individual differences in domain-general EFs to the process of DT, in terms of the serial order effect in both originality and the domain-specific processes of memory retrieval and mental operations. Studying the association of DT and EFs in young children is relevant for two reasons: (1) Research has suggested that formal schooling reduces DT because of the strong focus on convergent thinking in teaching (Gralewski et al., 2016; Krampen, 2012; Russ & Wallace, 2013). Therefore, it is important to know if and how EFs affect DT at an early age, before formal schooling begins. (2) Young children demonstrate similar involvement of domain-specific effortful DT processes in novel idea generation during a DT task as adults (Bai, Mulder, et al., 2021; Gilhooly et al., 2007). Yet, to the best of our knowledge, no study to date has examined the relationship between EF and domain-specific DT processes in children this young.

Our specific research aims were as follows. First, we aimed to investigate whether the serial order effect of originality, as reported in 4-year-olds in our previous study (Bai, Mulder, et al., 2021), was also present in 5- to 6-year-olds. If replicated, this would strengthen the evidence base that the serial order effect—a well-known effect occurring in DT in older children and adults—is already present in very young children. Second, we aimed to test if the two most important DT processes that we identified in our previous study in 4-year-olds, i.e., *mental operations* and *memory retrieval*, would also occur in 5- to 6-year-olds and would be related to originality in a similar way at this age. Specifically, we expected that memory retrieval would occur relatively frequently overall but would not relate to originality, whereas mental operations would occur much less frequently overall but would relate positively to originality.

Third, in addition to testing whether we could replicate findings from our previous study, we aimed to expand on the previous study by exploring how domain-general EFs and domain-specific processes play a role in the unfolding of original ideas. To this end, we systematically investigated the interactions between domain-general EFs, domain-specific DT processes (memory retrieval and mental operations), and use rank in predicting the originality of uses. We formulated two hypotheses regarding the moderating effects. First, we expected the process of mental operations to become increasingly important as a predictor of

the originality of a use the later in the sequence of all uses mentioned for a particular object. This hypothesis was established based on the findings in our previous study (Bai, Mulder, et al., 2021), which suggested a gradual shift from responses based on associative memory processes (comparatively less original) to responses based on controlled mental operations (comparatively original) during DT. Second, we expected the effect of individual differences in EFs, including selective attention, on originality to increase with use rank based on the theoretical consideration that a gradual shift from automatic associative responses to controlled executive processing would demand stronger involvement of domain-general EF capacities and, hence, would enlarge the effect of individual differences therein. Relatedly, we expected an interaction effect of EFs and the DT process mental operations on originality, presupposing that reliance on domain-specific mental operations in generating novel uses for a particular object would be facilitated by a higher level of domain-general EF abilities (and impeded by a lower level in EFs).

Materials and Methods

Participants

The present study is part of a longitudinal research project which investigates the development of DT in young children. 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). One hundred and seven children (49 boys and 58 girls; age at the first measurement wave ranged from 3.87 to 5.10 years of age, $M = 4.44$, $SD = 0.26$) with informed, active parental approvals from four typical primary schools in the Netherlands participated in this research project. The longitudinal project followed children for about 1.5 years, from the age of 4 years. DT was measured every 6 months at age 4, 4 to 5, 5, and 5 to 6 years. The current study used data from the last measurement wave, in which children were also given a battery of domain-general EF tasks.

Measures

Divergent Thinking

To assess children's DT, we used the AUT task (Bai, Mulder, et al., 2021; Guilford, 1967). Children were first instructed that they would see a number of images of common objects and that, for each object, they should think of as many different and unusual uses as possible. Subsequently, a real newspaper was presented to children and used as an example object. The experimenter first asked children what a newspaper could be used for to start the conversation, then gave three examples of unusual uses of the newspaper and explained how she came up with these uses (e.g., "you can fold the newspaper, then you have a hat"), and eventually encouraged children to give unusual uses and explain their thinking. Occasionally, if children did not fully understand the instruction, additional examples were given. The purpose of using this example object was to ensure that children understood the task. Next, the formal testing began, and six pictures of common objects, including a lunch box, a tire,

an umbrella, a pencil, a shovel, and a toothbrush, were presented to children one by one. For each object, children were asked to generate as many different and unusual uses as possible and to explain their thinking processes in relation to the generation of these uses. To encourage children to generate more uses and to report on their thinking processes, the task was conducted with interactive dialogues (Van Someren et al., 1994). Children were prompted to think of more uses (e.g., “What else can you use a basket for?”), to elaborate on a generated use (e.g., “How do you do that?”, “Can you tell me more about it?”), and to explain how they came up with a use (e.g., “How did you come up with this idea?”, “Have you done it before?”). All children were prompted, but children who were less expressive or whose utterances were more idiosyncratic and hard to grasp received more prompts. No strict time limit was set for the test considering that children might differ from each other regarding the time needed to fully express their thinking. The test took about 15–35 minutes, including the instructions and necessary breaks. All test sessions were video-recorded for later coding.

Assessing DT Ability. Although it is common to derive several DT ability measures from tasks such as the AUT (e.g., fluency, flexibility, and originality; see Runco & Acar, 2012), we focused specifically on originality in the present study, as originality is generally considered the core feature of divergent and creative thinking (J. K. Smith & Smith, 2017; Runco & Jaeger, 2012; Simonton, 2012; Weisberg, 2015), while fluency and flexibility are regarded as facilitating originality (Forthmann et al., 2018; Nijstad et al., 2010; Silvia et al., 2008). Following past research (Krampen, 2012; Van de Kamp et al., 2016), we first coded all mentioned uses per object based on the type of actions implicated in these uses. For instance, using a brick “to build a house” and “to build a bridge” were coded as the same type of action, i.e., “to build something” for the object “brick” (see Appendix 3.A for a list of action types per object). Subsequently, we scored the originality of the action types, per object, based on how often that action type was mentioned in the whole sample, using the following equation: Originality of an action type = $1 - (\text{The number of participants who mentioned one or more uses that implicated this action type} / \text{The total number of participants})$. Eventually, all uses received a score for originality based on the originality of the implicated action type. Note that the originality of uses had a nested data structure: the originality scores of the uses were nested in objects and the objects were nested in children. These data were used in the main analyses.

In addition, we also derived child-level fluency and originality scores for descriptive purposes. Fluency referred to the average number of uses children generated across objects. Originality was calculated following two steps: first the originality scores of action types per object were summed, then the sums per object were averaged per child.

Domain-Specific DT Processes. All video recordings of the AUT test sessions were verbatim transcribed. Content relevant information, including children’s use of signs and gestures, objects in the test environment that children referred to, whether the test was interrupted, and so on, were noted in the transcripts to facilitate correct understanding of

children's utterances. The transcripts were then divided into episodes, with each episode covering a coherent stretch of dialogues that was related to a single generated use. Subsequently, the coding scheme of Bai, Mulder, et al., (2021) was applied by the first author to assess the process of DT. For the purpose of the present study, we focused on the two key domain-specific DT processes as identified in the previous study, namely memory retrieval, which was found to be most frequently referred to by children, and mental operations, which was found to uniquely predict the originality of uses (Bai, Mulder, et al., 2021). For each episode, we determined whether children, in explaining their thinking, referred to memory retrieval (scored 1) or not (scored 0) and to mental operations (scored 1) or not (scored 0). Note that within a particular episode, children could refer to none, one, or both DT processes. Table 3.1 presents the definitions and a number of examples of these processes.

Table 3.1

Definitions and Examples of Coded Divergent Thinking (DT) Processes in the Alternative Uses Task (Bai, Mulder, et al., 2021)

DT process	Definition	Examples (C = child; T = experimenter)
1. Memory retrieval	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. 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. " 2a. T: "Have you done that before? " 2b. C: "Yes, I have done it once. "
2. Mental operations	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. " 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. "

To establish inter-coder reliability, 15% of the transcripts were double-coded by a trained second coder (a master's student in educational sciences). Cohen's kappa's showed moderate to substantial agreement between the two coders, $\kappa = .80$ for memory retrieval and $\kappa = .69$ for mental operations (see Appendix 3.B for an overview of the inter-coder reliability). Similar to originality, the data of the DT processes also had a nested structure: the occurrence of a DT process per mentioned use was nested in objects and the objects were nested in children. These data were used in the main analyses.

In addition, we calculated child-level scores of the DT processes for descriptive purposes. Child-level DT process scores were computed as the average frequencies of

occurrence of the processes memory retrieval and mental operations in the generation of uses across objects.

Executive Functions

Several measures were used to assess domain-general EFs. In addition to tests of inhibition, shifting, and working memory, we included a test of selective sustained attention as a general component of EF which is especially relevant at younger ages (Hendry et al., 2016; Veer et al., 2017).

Inhibition. Two tasks were used and programmed in E-prime (Psychological Software Tools, <http://www.pstnet.com>) to assess inhibition. The *Go/NoGo* task measures the extent to which children can inhibit a dominant motor response (Newman et al., 1985). The test is reported to have satisfactory psychometric quality, with a test-retest reliability of .65 (Nguyen et al., 2018). Children were instructed to press the space key when “go” stimuli (pictures of ten common objects, e.g., a book) were shown on the screen and not to press the space key when the “no-go” stimulus was shown (a picture of a dog). After introduction to the task and a practice session, children performed the test session in which the go and no-go stimuli were shown 40 and 20 times, respectively. Each stimulus was presented for 1000 ms, and the intervals between two stimuli varied from 200 to 1200 ms so that children were not able to predict the appearance of a stimulus. The order of stimulus presentation was pre-randomized and the same for all participants. The frequency at which children pressed the space key when they were shown the no-go stimulus, and thus failed to inhibit their dominant motor response, was calculated.

The *Animal Stroop* task measures the extent to which children can inhibit their reaction towards a salient feature and name a sub-dominant feature of a stimulus (Van der Ven et al., 2012). The test-retest correlations of three separate measures derived from this task over an 18-month interval in a sample of children aged 6 years at the first measurement occasion are reported to range between .28 and .33, and using confirmatory factor analysis all measures were found to load significantly on a combined inhibition-shifting factor (Van der Ven et al., 2012). In this task, children were first given the control block in which typical animal drawings of a cow, sheep, duck, and a pig were presented (head-body congruent stimuli). Children were instructed to give the name of the animal as soon as possible when a drawing appeared on the screen. In this block, each animal was presented 12 times, yielding a total of 48 trials. Next, children were given the test block in which a mix of drawings was presented: some drawings showed an animal body with a human head (control stimuli), and others showed an animal body with the head of another animal (head-body incongruent stimuli). Again, children were instructed to give the name of the animal for the body part (thus not for the more salient head part) as soon as possible each time a drawing appeared on the screen. There were 12 drawings for both the control and the head-body incongruent stimuli, and each drawing was presented four times, yielding a total of 96 trials. Presentation of each stimulus was preceded by a red fixation cross shown for 400 ms. The order of stimulus presentation

was pre-randomized and the same for all participants. Instructions and practice trials were given at the beginning of each block (control and test). During the practice trials, the experimenter gave feedback and corrected children's responses when they named the wrong animal (e.g., "Very good." or "Oops, that was a pig."), but during the test trials, no feedback was given. The percentage of head-body incongruent trials in which children correctly gave the name of the animal for the body part was calculated, following the recommendation of Van der Ven et al. (2012).

To strengthen the measurement quality, composite inhibition scores were computed by averaging the *z*-standardized scores of the Go/NoGo (reversed coded) and the Animal Stroop tasks (the correlation between the two measures is $-.21$; see Table 3.3). The composite scores were used for further analysis.

Shifting. Two tasks were used to assess shifting. The *Dimensional Change Card Sort* (DCCS) measures the extent to which children can shift to a new rule and abandon a previously correct rule. The DCCS has been widely applied to assess young children's cognitive flexibility, also referred to as shifting ability (Doebel & Zelazo 2015). The test-retest reliabilities of different versions of the DCCS are reported to range between .69 and .90 (Beck et al., 2011). An advanced version of the DCCS for 5-year-olds was used in the present study, which included three blocks (Beck et al., 2011; Zelazo, 2006). During the test, two boxes with a slot on the top side were placed next to each other in front of children, with one box labelled with a card that showed a blue star and the other box labelled with a card that showed a red truck. In the first block, children were instructed to sort six cards by color ("color game"), of which three showed a red star and three a blue truck. Therefore, the cards with a red star should go into the slot of the red-truck box, and the cards with a blue truck should go into the slot of the blue-star box. In the second block, children were instructed to sort a new set of the same cards by shape ("shape game"). Other than the color game, the cards with a red star should now go into the blue-star box, and the cards with a blue truck should go into the red-truck box. In the third block, children were instructed to sort a new set of twelve cards, of which six cards were identical to the cards used in the first and the second blocks, and another six cards had identical figures (three red-star cards and three blue-truck cards) but with a black border. The sorting rule was more complex in this block: if a card had a black border, the card needed to be sorted by color; if a card did not have a black border, the card needed to be sorted by shape ("advanced sort"). Every time a card was presented to the children, a rule reminder was given by the experimenter. The order of card presentation was pre-randomized and the same for all participants. Instructions were given at the beginning of each session. Additionally, following the instructions, for the color game, two practice trials were given; and for the shape game and the advanced block, the experimenter confirmed whether the children still remembered the sorting rules, for instance by asking "Can you show me where the stars go in the shape game?" and "What game do you play if there is a border?". Children's shifting scores were determined as follows: a score of 0 was given when children sorted less than five out of six cards correctly in the first block and thus

Chapter 3

failed on color sort; a score of 1 was given when children succeeded on color sort, but sorted less than five out of six cards correctly in the second block and thus failed on shape sort; a score of 2 was given when children succeeded on both color and shape sort, but sorted less than nine out of 12 cards correctly in the third block; finally, a score of 3 was given when children succeeded on both color and shape sort and sorted at least nine out of 12 cards correctly in the third block.

The *Animal Shifting* task measures the extent to which children can shift between applying different rules (Van der Ven et al., 2012). The test-retest correlations of measures derived from this task over an 18-month interval in a sample of children aged 6 years at the first measurement occasion are reported to be between .32 and .37, and all derived measures loaded significantly on a combined inhibition-shifting latent factor (Van der Ven et al., 2012). The task was programmed in E-prime. Children were first given a control block and then a shifting block, and each block consisted of 40 trials. In the control block, for each trial, a picture of an animal (cat, dog, bird, or fish) or a fruit (strawberry, pear, cherry, or banana) was shown centrally on the screen. Children were instructed to name the animal or fruit as quickly as possible. In the shifting block, for each trial, a picture of an animal and a picture of a fruit were shown simultaneously (one on the left side and one on the right side of the screen). Children were instructed to name either the animal or the fruit as quickly as possible. If the background color of the screen was yellow, children should name the fruit, and if the background color was purple, children should name the animal. Presentation of each stimulus was preceded by a red fixation cross shown for 700 ms. The order of stimulus presentation was pre-randomized and the same for all participants. Instructions and practice trials were given at the beginning of each block (control and shifting). During the practice trials, the experimenter gave feedback and corrected children's responses when they named the wrong animal or fruit (e.g., "Very good.", "Oops, that is not right. It was a pear."). Additionally, for the shifting practice trials, a rule reminder was given when children gave incorrect responses. No feedback was given during the test trials. The percentage of trials in which children correctly named the animal or fruit in the shifting block was computed as a measure of shifting ability, following the recommendation of Van der Ven et al. (2012).

To strengthen the measurement quality, composite shifting scores were computed by averaging the *z*-standardized scores of the DCCS and the Animal Shifting tasks (the correlation between the two tasks is .21; see Table 3.3). These composite scores were used for further analysis.

Working Memory. To measure children's working memory, a computerized verbal *Word Recall Backwards* task as adapted by Verhagen and Mulder (2013) was used. The task was programmed in E-prime. Children listened to series of two and three monosyllabic words voiced by a laptop computer and were instructed to recall the words out loud in the reversed order. After the instruction and a short practice session, children continued with the test, which included six trials: three two-word trials and three three-word trials. The order of

presentation of the trials was the same for all participants. The percentage of trials for which children correctly recalled the words in the reversed order was calculated as a working memory score and used in the analyses.

Selective Attention. To measure selective attention, a computerized visual search task programmed in E-prime for children aged 2 to 5 years was used (Mulder et al., 2014; Veer et al., 2017). In Veer et al. (2017), the test-retest correlation of this task administered at age 2.5 and 3 years using an age-appropriate version for toddlers was .55. In two studies, the selective attention measure was found to be significantly associated with other executive function measures, in particular spatial working memory, verbal short term memory, and inhibition (Mulder et al., 2014; Veer et al., 2017). In the present study, the version for 5-year-olds was used. In this test, children had to find and point to targets (elephants) as quickly as possible while ignoring distractors (bears and horses of similar size and color as the targets). Four test trials were given, presenting a grid-structured screen with targets and distractors. Each trial was shown for 40 s. The target to distractor ratio increased from 1:5 for the first two trials (presented on a 6×8 grid screen) to 1:8 for the third trial (presented on a 9×8 grid screen), and 1:24.5 for the fourth trial (presented on a 17×12 grid screen; the animal images in this display were smaller compared to previous trials to fit in the screen). Instructions and practice trials were given at the beginning of the task. After finishing each test trial, verbal and visual rewards were given (e.g., a picture of an elephant holding a flower was shown as a brief break in between test trials; the experimenter would say “Well done!”, “Good job!”). Additional instruction and two extra practice trials were given before the fourth test trial in order to familiarize children with the smaller images. Children were continuously prompted to keep searching and finding as many elephants as quickly as possible on each trial. Each time children pointed to a target, the experimenter would press a key so that the target would be crossed off by a diagonal blue line on the screen. The average number of targets children identified across all test trials was calculated and used as a measure of selective attention in the analyses.

Procedures

All children were tested individually in a separate room at their schools. Research assistants, students in educational sciences, were trained to administer the tests. They read the test instruction manual carefully, watched videos of example test sessions, conducted and videotaped a pilot test, and received extensive feedback. In addition, the first author supervised the first two to three test sessions of each assistant.

Data Analyses

Missing Data and the Final Sample

The current study included 102 out of 107 children of the first measurement wave of the longitudinal project (48 boys and 54 girls; age ranged from 5.45 to 6.53 years, $M = 5.93$, $SD = 0.27$). Four children had moved to other schools, and one child was very shy and refused

to participate in any test. Appendix 3.C shows an overview of the missing data and reasons for missingness across the different tasks.

Statistical Analysis

We investigated whether *use rank* (the rank number of a use in the sequence of all mentioned uses for a particular stimulus object), the occurrence of the domain-specific DT processes of *memory retrieval* and *mental operations*, domain-general *EFs*, and also the *interactions* between these variables predicted use originality. As the outcome variable use originality had a three-level nested data structure (i.e., Level 1: use, Level 2: object, and Level 3: child) and constituted proportional data, multilevel logistic regression analyses with a binomial distribution were run. As a first step, we tested the serial order effect of originality by building an intercept-only model (M1) and a model with the linear and quadratic effects of use rank (M2). As a second step, we examined whether the DT processes predicted the originality of uses. We built separate models to test the main effects of memory retrieval and mental operations (M3), considering that the power of the analysis might be limited because (a) the sample size of our study was rather small (102 children) and (b) the distributions of both DT processes were skewed. In particular, the DT process mental operations occurred rather infrequently. Next, to explore the role of domain-general EFs in DT, we added progressively the main effects of the EFs and several within-level and cross-level interaction effects to the model. Specifically, the following models were run, separately for the DT processes memory retrieval and mental operations: a model with the within-level interaction effect of DT process by use rank (M4); and for each EF separately, a model with the main effect of the EF concerned and the cross-level interaction effects of use rank by EF and DT process by EF, allowing the effects of use rank and DT process to vary across children (M5). The multilevel models in this study were fitted in SuperMix version 2.1 (Hedeker et al., 2008), using adaptive quadrature estimation with 20 quadrature points. All predictors were grand-mean centered before running the analyses to avoid multicollinearity. For interpreting the results, the $p < .05$ level of significance was applied.

Results

Descriptive Statistics

Descriptive statistics for the EF tasks and the aggregated child-level measures of the AUT are shown in Table 3.2. Children generated on average 4.21 uses per object on the AUT. For about 43% of the uses, children gave an explanation of their thinking processes by referring to memory retrieval, for 12% of the uses the explanation referred to mental operations, and for 5% of the uses children's explanations referred to both memory retrieval and mental operations. For the remaining uses (about 51%) either no explanation was provided or the given explanation could not be coded as one of the two DT processes. Table 3.3 shows the correlations between the EF measures and the child-level measures of the AUT, and Table 3.4 presents the number of uses by use rank regarding children's references to the DT processes of memory retrieval and mental operations.

Table 3.2

Descriptive Statistics of Executive Function (EF) and Child-Level Divergent Thinking (DT) Measures

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Skew.	Kurt.	Min.	Max.	
DT ability measures (mean scores per stimulus):								
Fluency	101	4.21	1.79	1.58	2.96	1.00	11.00	
Originality	101	2.13	1.29	1.93	5.43	0.09	7.90	
Frequency of DT processes (mean counts per stimulus):								
Memory retrieval	101	1.80	1.05	1.35	2.90	0	6.00	
Mental operations	101	0.50	0.69	2.59	7.49	0	3.67	
EF measures:								
Inhibition	Go/NoGo: number of errors	100	3.61	2.82	1.93	8.06	0	19.00
	Animal stroop: proportion accuracy	100	.93	0.08	−4.27	26.49	.38	1.00
Shifting	DCCS: level achieved	96	2.34	0.50	0.41	−1.21	1	3.00
	Animal shifting: proportion accuracy	102	.91	0.09	−2.07	5.06	.50	1.00
Working memory	Word recall backwards: proportion accuracy	98	.65	0.20	−0.001	−0.58	.17	1.00
Selective attention	Elephant test: average number of targets found	101	6.79	0.62	−0.82	0.92	4.75	8.00

Note. DCCS = The Dimensional Change Card Sort task.

Table 3.3

Correlations Between Age, Executive Functions, and Child-Level Divergent Thinking Measures

Variable	2	3	4	5	6	7	8	9	10	11
1. Fluency	.90**	.05	.67**	.44**	.04	.03	-.004	-.04	-.06	.08
2. Originality		.15	.52**	.59**	-.05	.10	.02	.01	-.04	.11
3. Age			-.05	.06	-.08	.10	-.06	.12	.06	.09
4. Memory retrieval	-.27**			.19 ⁺	.11	.04	-.09	-.08	.03	.11
5. Mental operations	.51***	-.16			-.15	.11	.10	.10	.01	-.05
6. Go/NoGo						-.21*	-.07	-.17	-.10	-.02
7. Animal stroop							.11	.38**	.07	.07
8. DCCS ^a								.21*	.19*	-.06
9. Animal shifting									.16	.21*
10. Word recall backwards										.14
11. Elephant task										

Note. DCCS = The Dimensional Change Card Sort task. Correlations below the diagonal are partial correlations corrected for fluency. Correlations in italics are nonparametric correlations and estimated with Kendall's tau_b, as DCCS consisted of ordinal data.

⁺ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$

Chapter 3

Table 3.4

Number of Uses and Counts of Divergent Thinking Processes by Use Rank

Use rank	<i>N</i> (uses)	Memory retrieval	Mental operations
1	598	212	12
2	497	260	54
3	420	206	66
4	334	151	61
5	232	93	32
6	147	54	25
7	96	41	19
8	64	21	9
9	45	17	6
10	35	12	4
11	24	7	3
12	19	7	4
13	12	4	2
14	8	1	2
15	5	1	1
16	3	0	1
17	3	0	1
18	3	1	0
19	2	2	0
20	1	1	0
21	1	0	0
22	1	0	0
<i>N</i> (uses)	2550	1091	302
<i>M</i> (use rank)	3.63	3.47	4.48
<i>N</i> (children)	101	99	75

Multilevel Regressions

The results of multilevel regressions are presented in Table 3.5 (models M1–M4) and Table 3.6 (models M5).

Serial Order Effect of Originality

First, we tested whether the serial order effect was present in the originality of uses. Compared to the intercept-only model (M1), adding the linear and quadratic effects of use rank (M2) significantly improved the model fit, as indicated by the change in model deviance, $\Delta_{\text{deviance}}(2) = 145.01, p < .001$. Both linear and quadratic effects of use rank were significant. The positive coefficient of the linear effect indicated that there was a serial order effect of originality, meaning that the later a use was generated in the sequence of all mentioned uses per object, the higher the probability that this use was scored as original. Conversely, the negative coefficient of the quadratic effect of use rank indicated that the increasing probability of uses being scored as original levelled off and even started to decrease towards the end of idea generation (see Figure 3.1 the black curve).

DT Processes and Originality

Next, we investigated whether memory retrieval and mental operations predicted the logit probability of the originality of uses in separate models (i.e., models M3-MR and M3-MO in Table 3.5 respectively). Adding the main effects of DT processes significantly improved the model fit compared to M2: for memory retrieval, $\Delta_{\text{deviance}}(1) = 5.82, p < .05$; for mental operations, $\Delta_{\text{deviance}}(1) = 20.97, p < .001$. The results showed that, although both memory retrieval and mental operations significantly and positively predicted the logit probability of the originality of uses, the effect of mental operations was somewhat larger than the effect of memory retrieval (for memory retrieval: odds ratio $[OR] = 1.13$; for mental operations: $OR = 1.38$).

Full Models With EFs for Memory Retrieval

Finally, we investigated how domain-general EFs, domain-specific DT processes, and use rank interacted in predicting the logit probability of the originality of uses. For the DT process memory retrieval, adding the within-level interaction of memory retrieval by use rank (see model M4-MR in Table 3.5) only marginally improved the model fit ($p < .10$), and the coefficient of the interaction was only borderline significant ($p = .07$). In Figure 3.1, we plotted the probability of originality by use rank for both conditions by whether references were made versus not made to the process of memory retrieval. As shown in this figure, the serial order effect of originality did not change much dependent on whether references were made to memory retrieval (the blue solid curve) or not (the blue dashed curve). Additionally, as shown in Table 3.6, in the full models (M5) which included also the main effect of EF and the cross-level interactions of use rank by EF and memory retrieval by EF¹, with each EF construct tested in a separate model, only the interaction of selective attention by use rank was borderline significant ($p = .052$).

Full Models With EFs for Mental Operations

For the DT process mental operations, adding the within-level interaction of mental operations by use rank significantly improved the model fit, and the coefficient of the interaction was significant (see model M4-MO in Table 3.5). This interaction effect is also plotted in Figure 3.1. As shown in this figure, for uses generated before the rank number of 10, which comprises about 97% of all mentioned uses, uses that were explained with references to the process of performing mental operations on the stimulus (the green solid curve) were evidently more original than uses that were explained without such references (the green dashed curve). For uses generated after the rank number of 10, which took up only

¹ Due to missing data on the EF measures, the number of participants included in the M5 models differed from the M4 models. As a consequence, a direct comparison of model deviances between M4 and M5 for each EF was not possible. Therefore, the change in model deviance is not reported.

Chapter 3

3% of all mentioned uses, an opposite pattern can be observed: uses that were explained with references to the process mental operations were less original than uses that were generated without such references. Next, largely in line with the findings for the process memory retrieval, in the full models that included the main effect of EF and the cross-level interactions of use rank by EF and mental operations by EF (M5 models in Table 3.6), only the interaction of selective attention by use rank was found to be significant. In Figure 3.1, we again plotted the probability of originality by use rank for high (the orange solid curve; one *SD* above the mean²) versus low selective attention (the orange dashed line; one *SD* below the mean). As this figure shows, before use rank three, there is no observable difference in the probability of originality of uses between both levels of selective attention, while beyond this rank children with lower selective attention generate consistently more original ideas. In addition, the difference between the levels of selective attention first increases, until approximately use rank 12, and then decreases gradually.

² The *SD* and mean were calculated based on the disaggregated data of selective attention at the use level.

Table 3.5

Multilevel Logistic Regression Analyses With the Originality of Uses as the Outcome Variable, Examining the Linear and Quadratic Effects of Use Rank (i.e., the Serial Order Effect) and the Effects of Divergent Thinking (DT) Processes in Predicting the Logit Probability of the Originality of Uses (N = 101, with 606 units at the object level and 2550 units at the use level)

Models	Intercept-model		Serial order effect		Memory retrieval (MR)				Mental operations (MO)			
	M1		M2		M3-MR		M4-MR		M3-MO		M4-MO	
Fixed part	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR
Intercept	-5.11 (0.03)***	0.006	-5.05 (0.03)***	0.006	-5.06 (0.03)***	0.006	-5.05 (0.03)***	0.006	-5.06 (0.03)***	0.006	-5.06 (0.03)***	0.006
Linear: Use rank			0.15 (0.01)***	1.17	0.16 (0.01)***	1.17	0.16 (0.01)***	1.17	0.15 (0.01)***	1.16	0.15 (0.01)***	1.16
Quadratic: Use rank			-0.014 (0.002)***	0.99	-0.013 (0.002)***	0.99	-0.014 (0.002)***	0.99	-0.013 (0.002)***	0.99	-0.012 (0.002)***	0.99
DT process					0.12 (0.05)*	1.13	0.14 (0.05)**	1.15	0.32 (0.07)***	1.38	0.39 (0.07)**	1.47
DT process × Use rank					-0.034 (0.02)*	0.97				-0.056 (0.03)*		0.95
Random part												
					Variance (s.e.)							
$\sigma^2_{\text{object level}}$	0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0001)	
$\sigma^2_{\text{child level}}$	0.00 (0.0002)		0.00 (0.0001)		0.00 (0.0001)		0.00 (0.0001)		0.00 (0.0001)		0.00 (0.0001)	
Deviance (<i>df</i>)	4386.80 (3)		4241.79 (5)		4235.97 (6)		4232.68 (7)		4220.82 (6)		4215.91 (7)	
Adviance (<i>Adv</i>)			145.01 (vs. M1, 2)***		5.82 (vs. M2, 1)*		3.29 (vs. M3-MR, 1)*		20.97 (vs. M2, 1)***		4.91 (vs. M3-MO, 1)*	
Scale ^a	0.56		0.50		0.50		0.50		0.49		0.49	

Note. Deviance = $-2 \times \log\text{-likelihood}$.

^a The estimations of *scale* indicates variance dispersion. In the current study, the scales are all smaller than 1, indicating that the observed variance of the outcome measure is smaller than the theoretical variance (i.e., $\pi^2/3$) of a binomial distribution.

* $p < .10$. ** $p < .01$. *** $p < .001$.

Table 3.6

Multilevel Logistic Regression Analyses With the Originality of Uses as the Outcome Variable, Examining Interaction Effects Between Executive Functions (EFs), Divergent Thinking (DT) Processes, and Use Rank in Predicting the Logit Probability of the Originality of Uses

Models	M5: Memory retrieval (M5-MR)						M5: Mental operations (M5-MO)					
	Inhibition		Shifting		Working memory		Selective attention		Inhibition		Shifting	
Fixed part	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR	Coeff. (s.e.)	OR
Intercept	-5.06 (0.03)***	0.006	-5.06 (0.03)***	0.006	-5.06 (0.03)***	0.006	-5.05 (0.03)***	0.006	-5.07 (0.03)***	0.006	-5.06 (0.03)***	0.006
Linear Use rank	0.15 (0.01)***	1.17	0.15 (0.01)***	1.17	0.15 (0.01)***	1.17	0.16 (0.01)***	1.17	0.15 (0.01)***	1.16	0.15 (0.01)***	1.16
Quadratic Use rank	-0.015 (0.002)***	0.99	-0.014 (0.002)***	0.99	-0.014 (0.002)***	0.99	-0.015 (0.002)***	0.99	-0.013 (0.002)***	0.99	-0.013 (0.002)***	0.99
DT process	0.15 (0.05)**	1.16	0.17 (0.05)**	1.18	0.15 (0.05)**	1.16	0.15 (0.05)**	1.16	0.38 (0.08)***	1.46	0.38 (0.08)***	1.47
DT process × Use rank	-0.031 (0.02)	0.97	-0.035 (0.02)*	0.97	-0.032 (0.02)	0.97	0.034 (0.02)*	0.97	-0.060 (0.03)*	0.94	-0.056 (0.03)*	0.95
EF	0.031 (0.04)	1.03	0.040 (0.03)	1.04	-0.014 (0.03)	0.99	0.015 (0.03)	1.01	0.011 (0.04)	1.01	0.023 (0.03)	1.02
Use rank × EF	0.019 (0.01)	1.02	0.009 (0.01)	1.01	0.003 (0.01)	1.00	-0.017 (0.00)*	0.98	0.019 (0.01)	1.02	0.009 (0.01)	1.01
DT process × EF	-0.039 (0.07)	0.96	-0.008 (0.07)	0.99	-0.007 (0.05)	0.99	-0.010 (0.05)	0.99	0.004 (0.11)	1.00	-0.013 (0.09)	0.99
Random part												
						<i>Variance (s.e.)</i>						
σ^2 object level	0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0002)		0.00 (0.0001)		0.00 (0.0001)	
σ^2 child level	0.0002 (0.0007)		0.0002 (0.0007)		0.0002 (0.0007)		0.0001 (0.0005)		0.0001 (0.0005)		0.0001 (0.0005)	
σ^2 use rank at child level	0.0003 (0.0004)		0.0003 (0.0004)		0.0002 (0.0004)		0.0001 (0.0003)		0.0003 (0.0004)		0.0003 (0.0004)	
σ^2 DT process at child level	0.00 (0.0004)		0.00 (0.0004)		0.00 (0.0005)		0.00 (0.0005)		0.0006 (0.0024)		0.0008 (0.0029)	
Deviance (df)	4011.66 (15)		3964.13 (15)		4120.28 (15)		4165.49 (15)		3997.16 (15)		3952.87 (15)	
Scale ^a	0.50		0.50		0.50		0.50		0.49		0.49	

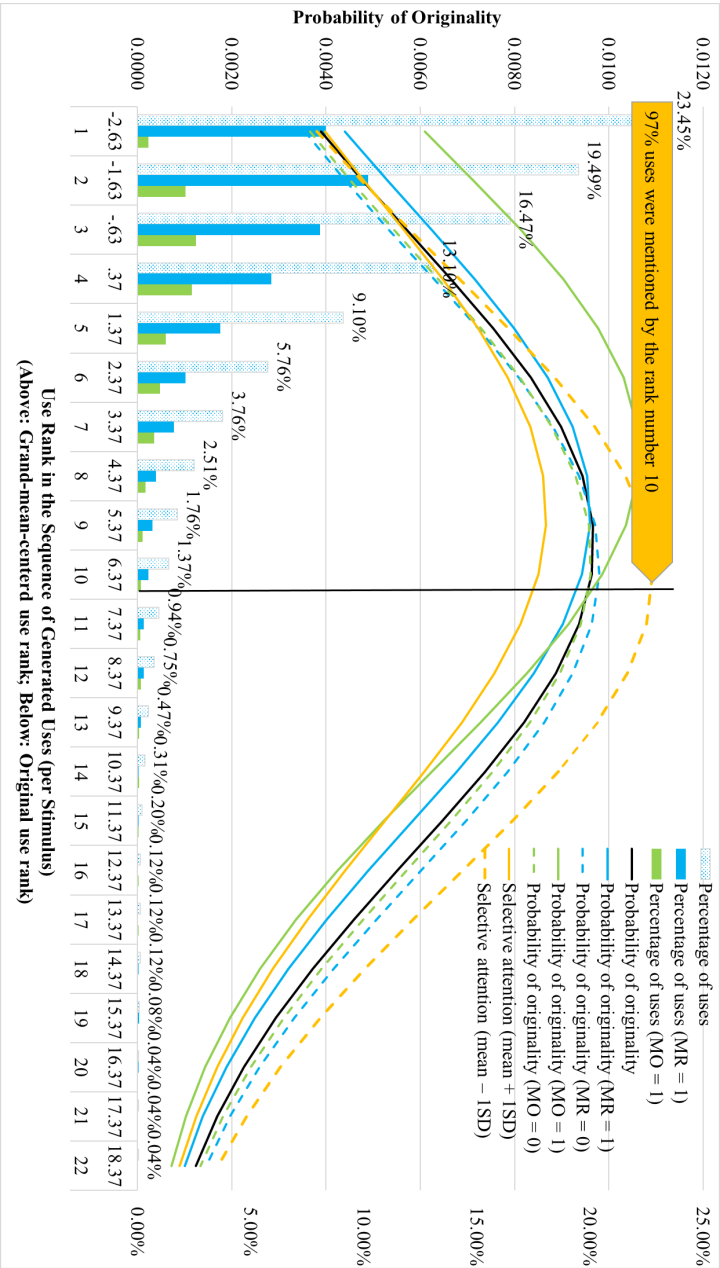
Note. Deviance = $-2 \times \log$ -likelihood. *N* is 97, 95, 97, and 100 for models including inhibition, shifting, working memory, and selective attention respectively.

^a The estimations of τ^2 scale^a indicates variance dispersion. In the current study, the scales are all smaller than 1, indicating that the observed variance of the outcome measure is smaller than the theoretical variance (i.e., $\pi^2/3$) of a binomial distribution.

⁺ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 3.1

The Probability of the Originality of Uses by Use Rank Under Different Conditions



Note. MR = Memory retrieval; MO = Mental operations. For a better visual presentation, we plotted this figure based on estimates of an additional full model ($N = 100$) which included the linear and quadratic effects of use rank (coefficients were 0.15 and -0.013 respectively), the main effects of memory retrieval, mental operations, and selective attention (coefficients were 0.14, 0.37, and 0.020 respectively), and interaction effects of use rank by memory retrieval, use rank by mental operations, and use rank by selective attention (coefficients were -0.028 , -0.054 , and -0.016 respectively). The estimations of coefficients were very close to results presented in Table 3.5 and 3.6 (difference ranges from 0.001 to 0.01).

Discussion

The goal of the present study was twofold. First, we aimed to replicate the findings of our previous study on DT in 4-year-olds (Bai, Mulder, et al., 2021). Specifically, we investigated whether the serial order effect of originality would occur again in the same group of children, but now at the age of 5 to 6 years. In addition, we investigated whether the two domain-specific DT processes, *retrieval from episodic and semantic long-term memory* and *performing mental operations on the stimulus*, predicted the originality of the generated ideas. The results largely confirmed our expectations: (1) There was a serial order effect regarding the originality of mentioned uses within the AUT, that is, uses generated later in the sequence of all mentioned uses of a particular stimulus object were more original, at least until, roughly, rank 10 in the sequence of mentioned uses (after 97% of all uses were mentioned), as Figure 3.1 reveals. (2) The DT processes memory retrieval and mental operations both positively predicted the overall originality of uses, while mental operations showed a stronger effect on originality than memory retrieval.

Second, the present study also aimed to deepen our understanding of how controlled executive processes operate during DT and how they may moderate the serial order effect of originality. We systematically investigated the interactions between domain-general EFs, domain-specific DT processes, and use rank, regarded as a quasi-time variable, in predicting the originality of uses. The findings can be summarized as follows. First, there was an interaction effect of the DT process mental operations by use rank in predicting the originality of uses. Second, selective attention, but none of the other EFs examined in this study (shifting, inhibition, and working memory) moderated the serial order effect of originality. Finally, there was no interaction effect between domain-general EFs and domain-specific DT processes on originality. Next, we discuss each of these findings in turn.

The serial order effect of originality in DT is already well established in older children and adults (Christensen et al., 1957; Mednick, 1962; Milgram & Rabkin, 1980; Wang et al., 2017; Ward, 1969). The present study among 5- to 6-year-old children replicates our previous study with the same children when they were 4 years of age (Bai, Mulder, et al., 2021), and both studies add to the body of evidence by confirming the serial order effect in young children. A further finding from the current study, replicating our previous work, was that the domain-specific DT process mental operations was a positive predictor of originality. We proposed, as an interpretation, that performing mental operations on the stimulus is executive in nature, given that during this type of processes, one would have to hold an image of the object in mind while manipulating it, inhibit the perceptual salience of the object in its usual whole integrated form, and shift attentional focus from global, holistic to local, singled-out features of the object in order to operate on these features. In particular, the mental process of singling-out properties or parts of whole objects and subsequently (re)combining these properties or parts into a new whole is considered conducive for generating original ideas in several theoretical (Lockman 2000) and empirical studies (Bai, Mulder, et al., 2021;

Forthmann et al., 2019; Gilhooly et al., 2007; Nusbaum & Silvia, 2011; Van Dijk et al., 2020). The present findings provide additional, though indirect, support for this interpretation.

The current study further explored how domain-specific processes (mental operations, memory retrieval) and domain-general executive processes (EFs) relate to the generation of original ideas and in particular to the serial order effect. We tested our hypotheses by examining the interactions of DT processes by use rank, DT processes by EFs, and EFs by use rank.

First, we expected the process of mental operations to become an increasingly important predictor of originality with increasing rank of the uses mentioned for a particular stimulus object, due to a gradual shift from responses based on associative memory processes (comparatively less original) to responses based on controlled mental operations (comparatively original). There was indeed an interaction effect of mental operations and use rank in predicting the originality of uses. However, contrary to our expectation, the interaction effect was negative. Based on the plot in Figure 3.1 (the green curves), the effect of mental operations on originality was already present in the initial stage of generating novel uses, remaining stable roughly until rank number six, but declining thereafter and turning even into a negative effect after rank 10 for the remaining 3% of all mentioned uses. In our previous study with the same children at age 4 years, based on a plot that modelled the occurrence probability of the process mental operations in relation to generated uses, the use of mental operations was virtually absent in the beginning of the sequence (Bai, Mulder, et al., 2021). A possible explanation is that the AUT was approached in a different way by the children in the current study. After successive testing with the AUT (also on two measurement occasions between 4 and 5- to 6 years; see Chapter 5) and due to being older and more accustomed to the structured context of the kindergarten, children may have understood the task demands better, complied more with the instruction of coming up with real novel uses, and therefore shifted almost immediately to the use of controlled executive processes, in particular mental operations, instead of initially relying on spontaneous memory-based associations.

To support this interpretation further, also the involvement of the DT process memory retrieval was different in the current compared to the previous study. Memory retrieval was considered to be mainly automatic and associative and to rely on the automatic spreading of activation in episodic and semantic long-term memory (Gilhooly et al., 2007; Mednick, 1962; Sowden et al., 2015). Therefore, memory retrieval was not expected to be a positive predictor of originality. The main effect of memory retrieval on originality, however, was positive and significant in the current study, though less strong than the effect of mental operations. In the previous study, there was no main effect of memory retrieval, and most importantly, memory retrieval was especially frequently referred to by the children explaining how they came up with an idea early in the sequence, right from the start, whereas mental operations were proportionally more frequently referred to in a later stage in the sequence of generated uses

(Bai, Mulder, et al., 2021). Both the positive main effects of the DT processes mental operations and memory retrieval, and the timing of these effects underscore that children may have approached the AUT differently at age 5 to 6 years compared to when they were 4 years of age. Note that also in the study of Gilhooly et al. (2007) among university students, memory-based DT processes positively predicted originality. Thus, together, the findings of these studies seem to suggest that very young children (i.e., 4-year-olds) may mostly rely on memory processes to generate relatively mundane ideas (e.g., “a basket is for carrying food, because I have seen that before”; Bai, Mulder, et al., 2021), while older children and adults appear to be better able to use—potentially unique—memories to generate original ideas (e.g., “[use a toothbrush] to make a giraffe...I have seen that in films, and they use cardboard and sometimes also toothbrushes and other things, and then you make an insect...I had seen once that they used the toothbrush[es] as the neck and the legs”). Clearly, further longitudinal work is required to test this hypothesis.

Finally, for both mental operations and memory retrieval, the interaction effect with use rank was negative, and the plots in Figure 3.1 revealed a turning point for both processes between ranks eight to 10 regarding the relation with originality, which turned negative. A possible explanation for this unexpected finding is that effectively applying mental operations and also memory retrieval requires effort, and children may have become tired after effortful thinking up about six or seven uses. If depletion indeed also played a role in using memory retrieval, this may further support the idea that also the use of this process differed from the previous study in that, instead of generating automatic, unfiltered responses from memory, additional control of the novelty of a generated use was applied by the children by inhibiting associative, yet less original responses.

Second, we hypothesized that individual differences in EFs would interact with the DT process mental operations in the prediction of originality. However, none of the main effects of the EFs shifting, inhibition, and working memory on originality (all tested in separate models), nor any of the interaction effects of these EFs and mental operations was found to be significant, which is inconsistent with previous research involving mostly older children and adults. Although the association between inhibition and DT measures such as fluency and originality is inconclusive based on previous research, shifting and working memory have been rather consistently found to be positively related to DT measures in most studies (De Dreu et al., 2012; Kharkhurin & Wei, 2015; Lee & Theriault, 2013; Lu et al., 2017; Nusbaum & Silvia, 2011; Yeh et al., 2015; Zabelina et al., 2019; Zabelina & Robinson, 2010). Possibly measurement issues have played a role here. There were ceiling effects in the two inhibition tasks, indicating that these tasks may have been too easy for children in the current age range, although the tasks were selected for being age-appropriate. Further research is needed to investigate if and how EFs and DT are related at this age.

Like the other domain-general EFs, selective attention had no main effect on originality, contrary to the expectation. However, a significant interaction effect of selective attention

and use rank on originality was observed, that is, selective attention was found to moderate the serial order effect of originality. This moderation effect, plotted in Figure 3.1 (the orange curves), however, was negative and indicated that children with a lower level of selective attention produced more original responses over time than children with a higher level of selective attention already, from about rank five onwards. This result is in line with the findings of a number of recent studies on the role of inhibition and selective attention, showing an advantage for creativity and DT of low inhibition and low selective attention (Radel et al., 2015; Stolte et al., 2019; Van Dijk et al., 2020; Zabelina, 2018). A possible explanation is that lower selective attention facilitates perceiving a broader set of external stimuli (such as irrelevant objects in the test environment or additional test stimuli to promote novel combinations of stimulus features; cf. Van Dijk et al., 2020) and promotes a broader spreading of activation to more remote and unrelated memory representations in semantic and episodic memory (Benedek, Könen, et al., 2012; Brown, 1973; Mednick, 1962), which would then lead to more original responses.

To summarize the findings, the present study provides tentative evidence for a complex interplay of different cognitive processes during DT, both controlled (the DT processes mental operations and memory retrieval) and less controlled (uncovered by the facilitating effect of low selective attention). The role of controlled processes early in the sequence of generating novel uses for a particular stimulus, as found in this study, seemed to reflect children's compliance with the task demands, while the uncontrolled processes, facilitated by a broadened attentional scope, gradually replaced the controlled processes, possibly due to depletion of the latter, towards the end of the sequence.

Strengths, Limitations, and Future Research

To the best of our knowledge, the present study is the first to systematically study how domain-general and domain-specific executive processes jointly operate during DT in young children with as yet limited formal schooling. Applying multilevel logistic regression analysis allowed us to study in a fine-grained way how DT upon a presented stimulus unfolds in real time and is moderated by both domain-general (selective attention) and domain-specific executive processes (mental operations, memory retrieval). In this way, the present study revealed a complex interplay of different cognitive processes underlying the generation of original responses in young children. A particular strength of the current approach was the multilevel modelling of both originality and DT processes. Detailing the processes of DT and creativity in future research may yield more conclusive evidence in particular regarding how individual differences in domain-general cognitive abilities such as EFs play out in the unfolding process of DT and creativity.

The present study suffered from a number of limitations. A clear limitation is the small sample size relative to the complex statistical models, which may have reduced the power of the analyses. The power problem may have been enlarged by another limitation, the ceiling effects in some of the EF tasks that were apparently too easy for children in the current age

range. Therefore, the fact that several expected effects were not found in the current study may not point to an absence of these effects but rather that these effects could not be detected. New research is recommended to test the current hypotheses again with larger samples, using EF tasks that are better able to capture the full ability range of children in the current age range.

Another notable limitation relates to the concept of EF and how to measure it in young children. In the current study, we treated inhibition, shifting, working memory, and selective attention as different executive functions and used separate tasks to measure them, following common practice (cf. Miyake et al., 2000). However, EFs may be less differentiated in young children than they are in school-age children and adults, and they may have a different underlying latent factor structure (Karr et al., 2018; Miyake & Friedman, 2012; Van der Ven et al., 2012). We recommend future research to carefully select age-appropriate EF tasks and to examine their latent factor structure prior to the main analysis which is feasible when working with larger samples than the one in the current study (for more elaborate recommendations, see Mulder et al., 2014; Van der Ven et al., 2012; Veer et al., 2017).

Other limitations also deserve attention in future research. Regarding the AUT task, the test environment and the procedure of giving verbal prompts should be more standardized in order to reduce random noise (e.g., objects present in the test environment could affect DT performance) and experimenter-variance (e.g., experimenters had to decide when and how many prompts were given to children). Finally, the current study was conducted in a sample of children attending kindergarten classrooms within Dutch primary schools. Given the wide variety of care and education arrangements in early childhood across the world, replication of the findings of the current study in different educational contexts is clearly warranted.

Conclusions

The present study replicated the findings of a previous study and confirmed that there is a serial order effect in the originality of ideas generated by young children on a widely used DT task, namely the AUT. Both domain-general EFs (selective attention) and domain-specific executive processes played a role in the real-time unfolding of original ideas, showing a complex interplay during DT. The processes of memory retrieval and applying mental operations on the stimulus were positively related to the early emergence of original ideas, while a lower level of selective attention facilitated original ideas in a later stage of idea generation. We could not confirm the expected role of the domain-general EFs shifting, inhibition, and working memory, as none of these EFs predicted originality or moderated the serial order effect in originality. The current study underscores the value of a process approach in studying DT. Future research is needed to further elucidate the processes that play a role in the generation of original ideas from early childhood to adulthood.

Chapter 4

The Development of Divergent Thinking in Four- to Six-Year-Old Children

This chapter is submitted for publication: Bai, H., Mulder, H., Moerbeek, M., Leseman, P. P. M., & Kroesbergen, E. H. (2021). The development of divergent thinking in four- to six-year-old children.

Author Contribution: *Conceptualization: H.B., E.K., P.L., and H.M.; methodology: H.B., E.K., P.L., H.M., and M.M.; investigation: H.B.; formal analysis: H.B., and M.M.; writing–original draft preparation: H.B.; writing–review and editing: E.K., P.L., H.M., and M.M.; supervision: E.K., P.L., and H.M.; funding acquisition: H.B., and E.K.*

The Development of Divergent Thinking in Four- to Six-Year-Old Children

Abstract

The present study investigated the development of divergent thinking (DT) in early childhood. A total of 107 4-year-old children were followed for 1.5 years. Children's DT was assessed with the Alternative Uses Task (AUT) every 6 months, four times in total. Within the AUT, children were not only asked to generate ideas for novel uses of common objects but also prompted to explain how they came up with these ideas. Based on the ideas children generated, two product measures were computed: fluency and originality. In addition, by coding children's explanations of their thinking processes, two process measures were computed: memory retrieval and mental operations. Multivariate multilevel regression analyses revealed substantial growth in fluency and originality with age. Increases with age were also observed in the number of references children made to the DT processes memory retrieval and mental operations when explaining their thinking processes. These findings contribute to the evidence base regarding the development of DT in early childhood and provide a first glance at the development of the specific thinking processes underlying young children's DT.

Keywords: divergent thinking, developmental trajectory, thinking process, young children, Alternative Uses Task, longitudinal research, multivariate multilevel analysis

Divergent thinking (DT) is an important and widely studied component of creativity (Runco & Acar, 2012). Compared to convergent thinking that usually follows logical (and largely linear) steps of thinking to reach a certain solution, DT is a thought process characterized by exploring many possible solutions to a problem (Guilford, 1956; Wang et al., 2017). DT is considered to be essential for innovation and creative problem solving (Bijvoet-van den Berg & Hoicka, 2014), particularly regarding ill-structured or open problems that require multifaceted exploration (Jonassen, 1997). Therefore, in contemporary society where rapidly changing technologies pose new challenges and increasing demands on inventiveness and creativity, fostering creativity and DT is increasingly recognized as an important educational objective in childhood (Kupers et al., 2019). Although it has recently been shown that the development of DT already starts in 1-year-olds (Hoicka et al., 2016), there is still a lack of knowledge about how DT develops in early childhood. Therefore, the present study aimed to investigate the developmental trajectory of DT in kindergarteners from 4 to 6 years of age. Specifically, in addition to studying the developmental trajectory of DT based on the product measures of fluency and originality, as is common in the field, we investigated the developmental trajectory of DT based on two specific thinking processes that have been found to underlie the generation of ideas during DT in children and adults: memory retrieval and mental operations (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021; Gilhooly et al., 2007).

Development of DT in Early Childhood: A Product Perspective

The few studies that have examined the development of DT in early childhood (Bijvoet-van den Berg & Hoicka, 2014; Daugherty, 1993; Krampen, 2012), all investigated DT from a product perspective. This means that children's DT ability was assessed based on outcome measures of DT tasks, yielding "product" performance measures such as: (1) *fluency*, i.e., the number of ideas generated by children, (2) *flexibility*, i.e., the number of conceptual categories occurring in the generated ideas of children, and (3) *originality*, i.e., the extent to which the generated ideas are statistically rare or rated as original (Reiter-Palmon et al., 2019).

While a temporary decline in DT (as measured with product measures) tends to occur when children enter primary school (Barbot et al., 2016; Krampen, 2012; G. J. W. Smith & Carlsson, 1983; Urban, 1991), previous studies mostly reported a steady increase in DT with age during the kindergarten years, when children do not yet receive formal education. In one study, for instance, Bijvoet-van den Berg and Hoicka (2014) tested DT in children aged 2 to 4 years with the Unusual Box Task (UBT). In the UBT, children are guided to play independently with a set of materials, including a novel box and several unconventional objects. By analyzing the actions that children demonstrated during exploring the unusual box, Bijvoet-van den Berg and Hoicka found that the number of unique actions displayed by children (i.e., fluency) and the originality of these actions were both positively correlated with age. In another study, Krampen (2012) found that children's fluency and flexibility as measured by six DT tasks increased over the kindergarten years, from age 4 to 6 years.

However, in one descriptive, small scale cross-sectional study, Daugherty (1993) reported that children's DT decreased from the age of 3 to 5 years and then started to increase from the age of 5 to 6 years.

Studying DT From a Process Perspective

In addition to studying DT based on product measures, an increasing number of recent studies have shed light on the thinking processes involved in the generation of ideas during DT (Beaty & Silvia, 2012; Forthmann et al., 2019; Gilhooly et al., 2007; Hass, 2017; Wang et al., 2017). In one study, using a detailed coding scheme of DT processes, Gilhooly and colleagues (2007) analyzed university students' think-aloud reports which were gathered while students were performing the Alternative Uses Task (AUT). The AUT is a widely applied DT test in which participants are asked to generate alternative uses for common objects (e.g., a brick; Guilford, 1967). They found that two thinking processes were conducive to DT: (a) the process of retrieving alternative uses from episodic and semantic long-term memory ("memory retrieval" from here on) and (b) the process of imagining disassembling the object in parts and recombining these parts in order to come up with new uses ("mental operations" from here on). More specifically, memory retrieval was observed to occur relatively early in the process of generating new uses of a particular stimulus object and the overall frequency of this process during the AUT was found to predict participants' fluency and originality. In contrast, mental operations occurred notably later in the process of generating new uses of a stimulus object, and its overall frequency during the AUT uniquely predicted participants' originality. In two previous cross-sectional studies, we examined the thinking processes underlying DT in young children (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021; see Chapters 2 and 3). In these studies, children were prompted to report on their thinking processes while performing the AUT. Analyses of children's thinking reports confirmed the involvement of similar processes in children's DT as has been reported for adults by Gilhooly and colleagues (2007). Furthermore, the results of these studies revealed that: (1) *Memory retrieval* was associated with fluency but not originality in 4-year-olds, and with fluency and originality in 6-year-olds. And (2) *mental operations* was associated with fluency and originality at both ages. In addition, references in children's explanations to memory retrieval were found to occur relatively frequently at the beginning of a series of uses generated for a stimulus object, but the frequency of occurrence decreased with the increasing rank of a use in the sequence of all mentioned uses for an object. The opposite trend applied to the references children made to mental operations, which occurred very infrequently initially, particularly in 4-year-olds, but increasingly frequently later in the sequence.

The Present Study

The present study aimed to investigate young children's development of DT from 4 to 6 years of age, from both a product and a process perspective. We employed a longitudinal design and followed a cohort of children for 1.5 years, from the time that these children

entered kindergarten (around 4 years of age) until shortly before they enrolled in formal education (around 6 years of age). Children's DT was measured with the AUT task, approximately every 6 months, four times in total.

First, we aimed to investigate the development of DT from a product point of view by studying the change in fluency and originality scores on the AUT over time. The longitudinal data on fluency and originality were analyzed with multivariate multilevel regression models. In accordance with previous studies, we expected growth in fluency and originality with age (Bijvoet-van den Berg & Hoicka, 2014; Krampen, 2012). However, we did not formulate a specific hypothesis regarding whether the growth would be simply linear or would accelerate or decelerate with age.

Next, we aimed to investigate the development of DT from a process point of view. To achieve this aim, we prompted children to report on their thinking processes while generating ideas on the AUT through interactive dialogues (Van Someren et al., 1994). The DT processes memory retrieval and mental operations were coded from children's explanations on how they generated each of the uses per stimulus object of the AUT. As with the product measures, the process measures of DT were analyzed using multivariate multilevel regressions. Previous studies have suggested that the application of the DT processes memory retrieval and mental operations underlies fluency and originality (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021; Gilhooly et al., 2007), and fluency and originality tend to increase with age in early childhood (Bijvoet-van den Berg & Hoicka, 2014; Krampen, 2012). Therefore, we hypothesized that we would observe an increasing occurrence of references to the DT processes memory retrieval and mental operations in children's explanations of how they came up with their ideas with age. Yet, given the scarcity of previous research, we did not develop specific hypotheses regarding the growth pattern.

Finally, we ran a series of control analyses to investigate whether repeatedly testing children across the 4- to 6-year age period and working with different stimuli sets on the AUT affected children's performance. First, we selected two equivalent sets of stimuli for the AUT and used them alternately at successive measurement waves. We investigated whether stimulus set affected performance, as further outlined below. Second, we randomly excluded half of the children from the test at the second measurement wave. We hypothesized that, if a testing effect was absent, the change in children's DT scores across the first, third, and fourth measurement waves would not be different between children who were measured at the second measurement wave and children who were not.

Method

Participants

One hundred and seven children (49 boys and 58 girls; age ranged from 3.87 to 5.10 years, $M = 4.44$, $SD = 0.26$) were recruited at the first measurement wave from four typical primary schools in the Netherlands and were followed for 1.5 years. In the Netherlands, the

Chapter 4

vast majority of children enroll in kindergarten education when they turn 4. Kindergarten education is part of the primary school system and comprises the first two of a total of eight grades. Parents gave informed consent prior to their child's participation in the study in either written or digital form via email. After the first measurement wave, there were three more measurement waves. The time between each two consecutive measurement waves was roughly 6 months. At the second measurement wave, only half of the children, randomly chosen, were tested; at the third and fourth measurement waves, all children were tested (see Table 4.1 for an overview of the sample per measurement wave).

Table 4.1

An Overview of the Sample Size, Missing Data, Inter-Coder Reliability of Divergent Thinking (DT) Processes, Additional Measures, and Related Studies per Measurement Wave

Wave	First wave (Oct. – Dec. 2016)	Second wave (Apr. – May 2017)	Third wave (Oct. – Dec. 2017)	Fourth wave (Apr. – June 2018)
AUT stimuli	Set 1	Set 2	Set 1	Set 2
<i>N</i>	84 (44 girls and 40 boys)	51 (27 girls and 24 boys)	102 (54 girls and 48 boys)	101 (53 girls and 48 boys)
Missing /excluded data	4 children did not finish the test; 3 children's tests were not recorded due to technical problems; 16 children's data were excluded as the experimenter who administered the tests did not prompt children to report on their thinking processes.	54 children were randomly chosen to not be tested at this wave; 1 child did not finish the test; 1 child was sick on test days.	3 children moved to other schools; 1 child was shy and refused to attend the test; 1 child's video recording was missing.	4 children moved to other schools; 1 child was shy and refused to attend the test. 1 child was not available on test days.
Kappa's	MR = .88; MO = .85	MR =.87; MO =.73	MR =.80; MO =.84	MR =.80; MO =.69
Other measure	TTCT	TTCT	TTCT	TTCT & EFs ^a
Related study	Bai, Mulder, et al., (2021)	--	--	Bai, Leseman, et al., (2021)

Note. AUT = Alternative Uses Task; MR = Memory retrieval; MO = Mental operations; TTCT = Torrance Test of Creative Thinking.

^aEFs = Executive functions. Six tasks were used to measure children's working memory, inhibition, shifting, and selective attention.

Measures

The AUT task was used to assess children's DT ability (Bai, Mulder, et al., 2021; Gilhooly et al., 2007; Guilford, 1967). The experimenter first explained to children that they would be shown pictures of several common objects and that they were required to think of as many different and unusual uses as possible for each object. Next, the experimenter

showed children a real newspaper as an example object to explain the purpose of the task. At this step, the experimenter (a) started the conversation by asking children what a newspaper could be used for, then (b) gave three unusual uses of the newspaper (occasionally additional example uses were given when children did not understand the instruction) and explained the thinking processes underlying the generation of these uses, e.g., “if you are going to paint, I have seen once that my mother used a newspaper to keep the table clean” or “you can fold the newspaper, then you have a hat”, and finally (c) asked children to give unusual uses of the newspaper and explain their thinking processes in a similar way. After this example item, the experimenter presented six formal test stimuli, one by one, in a randomized order for each child. Two sets of six stimuli were used alternately in successive measurement waves (see Table 4.1). For each object, children were instructed to generate as many different and unusual uses as possible and to report on their thinking processes. During this process, the experimenter gave ample verbal prompts to encourage children (a) to generate more ideas, e.g., “What else can you use a basket for?”, (b) to elaborate on their ideas, e.g., “How do you do that?”, “Can you tell me more about it?”, and (c) to explain their thinking processes, e.g., “How did you come up with this idea?”, “Have you done that before?” (Bai, Mulder, et al., 2021). The tests were conducted without a time limit as is common in research with young children to accommodate differences in speech rate and expressive abilities. The whole test took between 15 to 35 minutes, including the instructions and breaks. All AUT test sessions were video-recorded and subsequently transcribed verbatim for further coding and scoring. Context information related to the test situation but not reflected in the verbalizations (e.g., children’s use of signs and gestures, interruptions of the test, et cetera) was additionally noted in the transcripts to facilitate the understanding and correct interpretation of children’s utterances. The transcripts served as the main source for the assessment of children’s DT; however, when doubts or questions arose during coding, video recordings of the test sessions were checked.

Product Measures

Two product measures, *fluency* and *originality*, were computed based on the uses that children generated for the given objects during the AUT.

Fluency. Fluency scores were obtained in two steps: (1) For each stimulus, the fluency score was the total number of distinct uses children generated for that stimulus. Two uses was considered to be distinct if either the involved actions or the mentioned objects that afforded the actions were different (Bai, Mulder, et al., 2021). For instance, using a brick “to build a house” and “to build a bridge” are distinct from each other. (2) The fluency scores of all stimuli were averaged per child.

Originality. Originality scores were obtained in three steps: (1) The generated uses were categorized based on the involved type of actions, for each stimulus separately (Bai, Mulder, et al., 2021). In this regard, using a brick “to build a house” and “to build a bridge” were considered as involving the same action type, i.e., “to build something”. (2) Each action type

Chapter 4

was scored on originality based on how frequently participants mentioned this action type, applying the following equation: Originality of an action type = $1 - (\text{The number of participants who mentioned one or more uses that implicated this action type} / \text{The total number of participants})$. (3) The originality scores for the action types implicated in a child's generated uses were summed per stimulus, and subsequently, the originality scores of all stimuli were averaged for each child.

Note that the categorization of uses and the calculation of the originality of action types were carried out for each object across waves. In other words, uses generated on the same set of stimuli but at different waves, that is, uses of the first and third measurement waves for the first stimulus set and uses of the second and fourth measurement waves for the second stimulus set respectively, were merged for the categorization of uses and the calculation of originality. This was done in order to allow comparison of the data of the first and third waves (1st stimulus set) and of the second and fourth waves (2nd stimulus set), respectively. Note that a separate control analysis was carried out to investigate the effect of stimulus set, as reported in the Result section. In addition, participants at different measurement waves were seen as independent units in the calculation of the number of participants for the originality computation given above. For example, a child who had generated the same use for an object at both the first and third measurement waves would contribute a count of two in calculating the number of participants who mentioned the related action type. This rule also applied to the calculation of the total number of participants.

Process Measures

Two process measures were computed based on the coding of children's explanations regarding how uses were generated. For this purpose, the coding scheme that was previously developed to analyze the process of DT in young children was applied (Bai, Mulder, et al., 2021). In particular, we focused on two DT processes which were found to predict fluency and originality in our previous studies (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021): (a) *memory retrieval*, this process was coded when children referred to retrieval from semantic or episodic memory, for example when referring to past experiences to explain a use, e.g., "I do that (use a toothbrush to brush teeth) everyday"; and (b) *mental operations*, this process was coded when children referred to performing mental operations on the stimulus while generating a use, e.g., "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."

The AUT transcripts were first divided into episodes. Each episode contained a use and the related explanations of this use as given by the children. Next, each of these episodes was coded for the occurrence of memory retrieval and mental operations: for each DT process, 1 was coded if children had made one or more references to this process, and 0 was coded if children had not made any reference to a process. Children could refer to none, one, or both DT processes in explaining how they came up with one particular use within one episode. Finally, the frequencies of occurrence of each DT process were averaged per child.

Prior to the formal coding, inter-coder reliabilities of both DT processes were established by having a second trained coder coded between 10% to 20% of the transcripts of each measurement wave. Cohen's Kappa's showed moderate to almost perfect agreement and are presented in Table 4.1.

Procedures

The present study is part of a longitudinal research project which was approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences of Utrecht University in 2016 (reference number: FETC16-066). In the entire longitudinal project several other measurements were used, including a figural creativity test (at all measurement waves) and a battery of executive function tests (only at the last measurement wave). During these tests, children were not required to report on their thinking processes. Therefore, the present study utilized only data of the AUT test in order to simultaneously investigate the development of the product and the process of DT. Other studies of this project have investigated the process of DT cross-sectionally (Bai, Mulder, et al., 2021) and examined how executive processes operate during DT (Bai, Leseman, et al., 2021) in young children.

Children were tested individually in a separate room at their schools. As briefly mentioned above, two equivalent sets of stimuli were used for the AUT test alternately at the four consecutive measurement waves in order to minimize testing effects across waves. The first set of stimuli included pictures of a hand towel, a brick, a fishnet, a basket, a broom, and a spoon and were used at the first and third measurement waves; and the second set of stimuli included pictures of a lunch box, a tire, an umbrella, a pencil, a shovel, and a toothbrush and were used at the second and fourth measurement waves. In addition, half of the children, randomly selected, were not tested at the second measurement wave in order to enable us to examine whether a testing effect was present as a consequence of repeated testing.

Data collection was carried out by trained bachelor's and master's students in educational sciences. The training included carefully reading the test manual, watching example test sessions and reading related transcripts, and practicing the test procedure prior to the actual data collection. In addition, in order to monitor the quality of test administration, the first author (a) supervised the first two to three test sessions in the field conducted by students and gave feedback when needed and (b) was present at test sessions of all students on a number of occasions across the data collection period.

Data Analysis

Sample Size and Missing Data

There were a number of children who were not tested or whose data had to be excluded from the analyses for different reasons. The sample sizes of each measurement wave and the reasons for missing or excluded data are shown in Table 4.1.

Analysis Plan

Multilevel regression analyses were applied to analyze the longitudinal data (level-1: study wave; level-2: child). In particular, children's age (in years) at each measurement wave was used as a time-varying (level-1) predictor of children's DT measures, given the wide age range at each measurement wave (Hox et al., 2017, p. 84). Specifically, we ran two series of *multivariate* multilevel regression analyses: (1) Fluency and originality as two theoretically related outcome measures of DT products were analyzed together in multilevel regression models. Likewise, (2) memory retrieval and mental operations as two related outcome measures of DT processes were analyzed together in multilevel regression models.

The analyses consisted of two parts. First, we examined whether a testing effect was present. This was done by testing whether being repeatedly tested at the second measurement wave influenced the change in children's DT scores across the other measurement waves. To this end, the change in children's DT over time was modelled with data of the first, third, and fourth measurement waves, and the *group* variable that defined if children were tested at the second measurement wave (1 = *tested at the second wave* and 0 = *not tested at the second wave*) was used as a level-2 predictor. Two models were run: (a) an intercept-only model and (b) a model with the main effects of age (level-1) and group (level-2) and the interaction effect of age by group (cross-level). In the second model, the effect of age was allowed to vary across children.

Next, we investigated the change in children's DT scores over time using data of all measurement waves. Three models were run: (a) M0, an intercept-only model; (b) M1, a model with linear and quadratic effects of age (level-1), allowing these effects to vary across children; (c) M2, a more parsimonious model of model M1, which excluded the quadratic effect of age and the random term of this effect for the product measures and excluded only the random term of the quadratic effect of age for the process measures (reasons why effects were removed or retained are explained below in the Results section); and (d) M3, a model with a dummy variable of *stimulus set* as a level-1 predictor, in order to examine the influence of stimulus set on children's change in DT over time.

All models were fitted in SuperMix version 2.1 (Hedeker et al., 2008) and the maximum likelihood estimation was used. Age was entered in all models grand-mean centered to avoid multicollinearity between the linear and quadratic terms. We used an alpha level of .05 for all statistical tests.

Results

Descriptive Statistics

The descriptive statistics as presented in Table 4.2 showed two patterns: (1) The means of all measures increased across measurement waves. Similarly, (2) the standard deviations of all measures increased across waves, indicating that individual differences on these measures gradually enlarged when children grew older.

Next, as shown in Tables 4.3, the correlations of DT measures between adjacent measurement waves were generally higher than .40 and were mostly significant, indicating fair stability. In addition, in Table 4.4 correlations of DT measures within each measurement wave are presented, showing that: (1) The two product measures fluency and originality were strongly inter-correlated (r 's were between .76 to .90 per wave), whereas the process measures were much less strongly inter-correlated (r 's were between $-.11$ to $.19$ per wave). And (2) both product measures were quite strongly correlated with each of the process measures, except for the second wave, where neither fluency nor originality was correlated with memory retrieval.

Table 4.2

Descriptive Statistics of Divergent Thinking (DT) Measures and Age per Measurement Wave

DT measure	Wave ^a	<i>M</i>	<i>SD</i>	Min.	Max.	Skew.	Kurt.
Fluency	1	2.67	1.25	0.50	8.50	1.58	4.83
	2	3.05	1.20	1.50	7.50	1.37	2.76
	3	4.14	1.65	1.50	11.33	1.11	2.74
	4	4.21	1.79	1.00	11.00	1.58	2.96
Originality	1	0.85	0.46	0.18	2.44	1.14	1.77
	2	1.38	0.87	0.11	4.37	1.50	2.61
	3	1.88	1.03	0.21	4.55	0.62	-0.17
	4	2.20	1.30	0.11	8.00	1.93	5.41
Memory retrieval	1	0.97	0.60	0.00	2.50	0.51	-0.37
	2	0.75	0.54	0.00	2.33	0.83	0.52
	3	1.40	0.74	0.17	3.83	0.87	0.71
	4	1.80	1.05	0.00	6.00	1.35	2.93
Mental operations	1	0.08	0.19	0.00	1.33	4.02	21.76
	2	0.25	0.42	0.00	2.17	2.62	8.63
	3	0.38	0.57	0.00	2.50	2.26	4.89
	4	0.50	0.69	0.00	3.67	2.59	7.49
Age	1	4.44	0.26	4.02	5.10	0.32	-0.92
	2	4.86	0.26	4.35	5.36	0.14	-0.93
	3	5.45	0.26	4.91	5.99	0.20	-0.83
	4	5.93	0.27	5.45	6.53	0.28	-0.80

^a *N* for wave 1-4 were 84, 51, 102, and 101 respectively.

Chapter 4

Table 4.3

Correlations of Each Divergent Thinking Measure Between Measurement Waves

Wave	W1	W2	W3	W4
Fluency above the diagonal; Originality below the diagonal.				
W1		.39*	.45**	.22 ⁺
W2	.48**		.33*	.44**
W3	.43**	.41**		.28**
W4	.20 ⁺	.45**	.40**	
Memory retrieval above the diagonal; Mental operations below the diagonal				
W1		.27 ⁺	.32**	.19 ⁺
W2	.46**		.09	.19
W3	.09	.26 ⁺		.40**
W4	.22 ⁺	.60**	.40**	

⁺ $p < .10$. * $p < .05$. ** $p < .01$.

Table 4.4

Correlations of Divergent Thinking Measures and Age Within Each Measurement Wave

Variable	V1	V2	V3	V4	V5
First measurement wave: $N = 84$, above the diagonal; Third measurement wave: $N = 102$, below the diagonal.					
V1. Fluency		.76**	.31**	.45**	.15
V2. Originality	.81**		.22*	.61**	.22*
V3. Memory retrieval	.29**	.31**		.16	.09
V4. Mental operations	.52**	.67**			.01
V5. Age	.16	.21*	.02	.25**	
Second measurement wave: $N = 51$, above the diagonal; Fourth measurement wave: $N = 101$, below the diagonal.					
V1. Fluency		.89**	-.03	.59**	.02
V2. Originality	.90**		-.005	.69**	.13
V3. Memory retrieval	.67**	.52**		-.11	.14
V4. Mental operations	.44**	.59**	.19 ⁺		.08
V5. Age	.05	.14	-.05	.06	

⁺ $p < .10$. * $p < .05$. ** $p < .01$.

The Testing Effect

To examine whether a testing effect was present, the change in DT over the first, third, and fourth measurement waves were compared between children who were measured at the second measurement wave and children who were not. Results of multivariate models were

similar for both the product and the process measures. First, adding the main effects of age and group and the interaction effect of age by group significantly improved the model fit compared to the intercept-only models: for product measures, $\Delta_{\text{deviance}}(13) = 138.95, p < .001$; for process measures, $\Delta_{\text{deviance}}(13) = 124.57, p < .001$. Second, the effect of age was significant on all measures: fluency, *Coeff. (s.e.)* = 1.04 (0.18), $p < .001$; originality, *Coeff. (s.e.)* = 0.89 (0.18), $p < .001$; memory retrieval, *Coeff. (s.e.)* = 0.54 (0.10), $p < .001$, and mental operations, *Coeff. (s.e.)* = 0.21 (0.07), $p < .01$. Third, the main effect of group and the interaction effect of age by group were not significant for any of the outcome measures (see Appendix 4.A for a full presentation of the model results). Thus, whether or not children were tested at the second measurement wave did not influence children's DT performance over time, indicating that a testing effect was absent. For conciseness, we did not report the results of the more complex models that included also the quadratic effect of age and the interaction effect of quadratic age by group, as also in those models, none of the effects related to group was significant.

Change in DT Over Time and the Influence of Stimulus Set

To investigate the developmental trajectory of DT, the change in children's DT over time was modelled using data of all measurement waves. In addition, the influence of stimulus set on the development of DT was examined. The results are presented in Table 4.5.

Product Measures

Compared to the multivariate intercept-only model (see M0 for product measures in Table 4.5), adding the linear and quadratic effects of age to the model and allowing these effects to vary across children (M1) significantly improved the model fit. Given that, in this model, the quadratic effects were not significant, nor were the variances as estimated for these quadratic effects, we ran a more parsimonious model without these effects (M2). The parsimonious model was favored because the deviance of the model was only slightly higher while the number of parameters decreased substantially, $\Delta_{\text{deviance}}(12) = 8.88, p > .05$. The results of the parsimonious model (M2) showed that (a) the linear effects of age on fluency and originality were significant, and (b) these effects did not vary across children. Next, adding stimulus set as a predictor (M3) improved the model fit, $\Delta_{\text{deviance}}(2) = 10.77, p < .01$. Nonetheless, while the linear effect of age on both fluency and originality was taken into account, the effect of stimulus set was only borderline significant on fluency ($p = .07$) and not significant on originality, indicating that stimulus set had no or only a very limited influence on children's development on product DT measures.

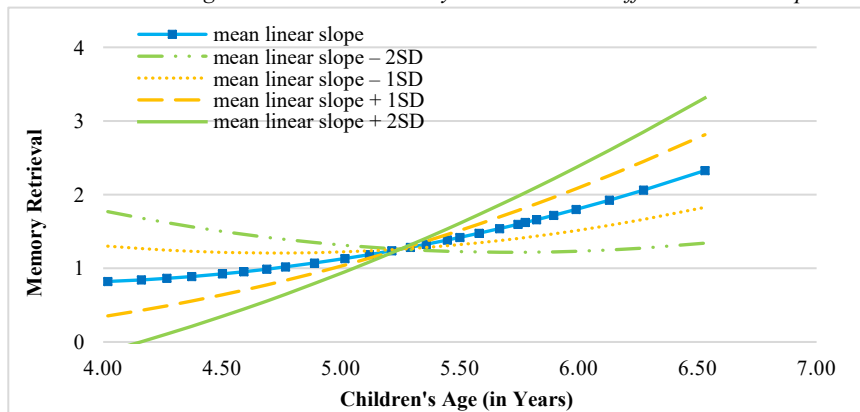
Process Measures

Compared to the multivariate intercept-only model (see M0 for process measures in Table 4.5), adding the linear and quadratic effects of age and allowing these effects to vary across children (M1) significantly improved the model fit. In this model, the quadratic effect of age was borderline significant ($p = .06$) on memory retrieval but not significant on mental

operations, and the variances of both quadratic effects were not significant. Therefore, we ran a more parsimonious model without the random term of the quadratic effects of age (M2). The more parsimonious model was favored given that the deviance of the model was smaller with less parameters, $\Delta_{\text{deviance}}(11) = -1.51, p > .05$. The results of the parsimonious model (M2) showed that (a) the linear effects of age on memory retrieval and mental operations were significant, (b) the quadratic effect of age was significant on memory retrieval but not on mental operations, and (c) the variance of the linear effect of age was significant on memory retrieval but not on mental operations. These results indicated that with increasing age children increasingly referred to both memory retrieval and mental operations in explaining how they came up with their ideas. In particular, the growth rate of referring to the process memory retrieval accelerated with age and varied across children. Based on the estimated linear slope of age on memory retrieval and its variance across children, we further computed a 95% predictive interval for the linear slope, which turned out to be between -0.19 and 1.37^3 . In Figure 4.1 we plotted the change in the measure of memory retrieval by age with different linear slopes. As shown in this figure, regardless of the overall accelerating trend (i.e., the blue solid curve), there was still a small proportion of the children who might decrease or stay at the same level over time with respect to memory retrieval (e.g., the orange and green dashed curves). Finally, adding stimulus set (M3) did not improve the model fit. While the linear and quadratic effects of age were taken into account, the effects of stimulus set on memory retrieval and mental operations were not significant, indicating that stimulus set had no influence on children's development in process DT measures.

Figure 4.1

The Estimated Changes Over Time in Memory Retrieval With Different Linear Slopes



³ In accordance with Hox et al. (2017, p.16), the varying regression slopes are assumed to follow a normal distribution with the mean equaling the estimated slope (i.e., 0.59 in this study) and the standard deviation equaling the square root of the estimated variances of the slopes (i.e., 0.39). Thus, in the current case, 95% ($z = 2$) of the random linear slopes of age was expected to lie between $(0.59 - 0.39 \times 2) = -0.19$ and $(0.59 + 0.39 \times 2) = 1.37$.

Table 4.5

Multivariate Multilevel Regressions: Examining the Changes in Divergent Thinking Measures Over Time and the Effect of Stimulus Set (N = 106, with 338 units at the measurement level)

Model	M0: Intercept-only		M1: Age		M2: Age		M3: Stimulus set	
	FLU	ORG	FLU	ORG	FLU	ORG	FLU	ORG
Fixed part	<i>Coeff. (s.e.)</i>							
Intercept	3.65 (0.11)***	1.65 (0.07)***	3.74 (0.12)***	1.64 (0.08)***	3.64 (0.11)***	1.64 (0.07)***	3.48 (0.14)***	1.66 (0.09)***
Linear age			1.02 (0.14)***	0.86 (0.10)***	1.01 (0.12)***	0.85 (0.08)***	1.13 (0.13)***	0.84 (0.09)***
Quadratic age			-0.28 (0.20)	-0.009 (0.12)				
Stimulus set							0.28 (0.16) ⁺	-0.02 (0.09)
Random slope								
$\sigma^2_{\text{linear age}}$			0.55 (0.34)	0.37 (0.14)*	0.16 (0.23)	0.15 (0.10)	0.18 (0.23)	0.15 (0.10)
$\sigma^2_{\text{quadratic age}}$			0.12 (0.58)	0.10 (0.23)				
Random effect								
σ^2_{child}	0.70 (0.20)***	0.27 (0.09)**	0.55 (0.25)*	0.26 (0.10)*	0.81 (0.19)***	0.36 (.08)***	0.81 (0.18)***	0.36 (0.08)***
σ^2_{wave}	2.14 (0.20)***	0.99 (0.09)***	1.57 (0.20)***	0.55 (0.07)***	1.54 (0.17)***	0.55 (0.06)***	1.51 (0.17)***	0.55 (0.06)***
Deviance (df)	1862.67 (8)		1702.97 (30)		1711.85 (17)		1701.08 (19)	
Δ deviance (Δ df)			159.7 (22)***, vs. M0		150.82 (9)***, vs. M0		10.77 (2)**, vs. M1	
Models	MR	MO	MR	MO	MR	MO	MR	MO
Fixed part	<i>Coeff. (s.e.)</i>							
Intercept	1.32 (0.06)***	0.32 (0.04)***	1.26 (0.07)***	0.30 (0.05)***	1.26 (0.07)***	0.30 (0.05)***	1.18 (0.09)***	0.31 (0.06)***
Linear age			0.61 (0.08)***	0.28 (0.05)***	0.59 (0.07)***	0.27 (0.05)***	0.64 (0.08)***	0.27 (0.05)***
Quadratic age			0.19 (0.10) ⁺	0.04 (0.07)	0.19 (0.10)*	0.04 (0.06)	0.23 (0.10)*	0.03 (0.06)
Stimulus set							0.12 (0.08)	-0.01 (0.05)
Random slope								
$\sigma^2_{\text{linear age}}$			0.23 (0.10)*	0.04 (0.04)	0.15 (0.07)*	0.02 (0.03)	0.16 (0.08)*	0.02 (0.03)
$\sigma^2_{\text{quadratic age}}$			0.04 (0.16)	0.01 (0.07)				
Random effect								
σ^2_{child}	0.13 (0.05)**	0.08 (0.02)***	0.15 (0.07)*	0.06 (0.03) ⁺	0.23 (0.05)***	0.10 (0.02)***	0.24 (0.05)***	0.10 (0.02)***
σ^2_{wave}	0.64 (0.06)***	0.22 (0.02)***	0.40 (0.05)***	0.18 (0.02)***	0.39 (0.04)***	0.17 (0.02)***	0.39 (0.04)***	0.17 (0.02)***
Deviance (df)	1363.54 (8)		1209.45 (30)		1207.94 (19)		1205.75 (21)	
Δ deviance (Δ df)			154.32 (22)***, vs. M0		155.60 (11)***, vs. M0		2.19 (2), vs. M1	

Note. FLU = Fluency; ORG = Originality; MR = Memory retrieval; MO = Mental operations.

⁺ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Discussion

The present study investigated young children's development of DT from the age of 4 to 6 years using a longitudinal design. A cohort of 107 children was followed during a period of about 1.5 years, from the time that these children enrolled in primary school kindergarten classrooms until shortly before they started with formal primary education. DT was assessed with the AUT task, and fluency and originality scores were derived from their answers. In addition, two thinking processes underlying DT were coded from children's explanations of how they generated each of the uses on the AUT: memory retrieval and mental operations. The analyses revealed several findings. First, children's DT improved with age: for fluency, originality, and mental operations, a linear increase was found, on average, and this increase was not found to vary across children; for memory retrieval, the increase accelerated with age, on average, and there was significant variance in this increase between children. Next, we found no evidence for significant effects on outcome measures of (a) using different AUT stimulus sets or (b) repeatedly testing children with the AUT. Below we discuss the findings regarding the product and process measures in more detail.

Development of DT Based on Product Measures

Past research on DT development mostly reported a steady increase with age during early childhood, based on product measures (i.e., fluency, flexibility, and originality; Bijvoet-van den Berg & Hoicka, 2014; Gralewski et al., 2017; Krampen, 2012). In line with these previous findings, we found that children's fluency and originality scores substantially increased over time from the age of 4 to 6 years. In particular, the results of the present study showed that these increases were linear, which, to the best of our knowledge, has not been previously investigated. On average, children improved from generating 2.67 uses at the first measurement wave ($M_{\text{age}} = 4.44$ years) to generating 4.21 uses at the fourth measurement wave ($M_{\text{age}} = 5.93$ years), with an effect size Cohen's $d = 1.00$. Performance at the fourth measurement was relatively similar to what has been reported in school-age children (Van Dijk et al., 2020) and adults (Gilhooly et al., 2007), who were found to give on average four to six uses per object within the AUT, although under different test conditions. In Van Dijk et al. (2020), fifth-grade children were presented with real objects and received limited verbal prompting; in Gilhooly et al. (2007), adult subjects were presented with object-words and not prompted at all. Possibly, the active prompting as applied in the present study has encouraged children to spend more effort, thereby promoting their performance on the AUT to a level comparable to older children and adults (Gilhooly et al., 2007; Van Dijk et al., 2020).

Development of DT Based on Process Measures

The results of the present study confirmed our hypothesis that children would also show an increase with age in the reported application of two key DT processes presumed to underlie the generation of novel uses in a DT task like the AUT, i.e., memory retrieval and mental operations (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021; Gilhooly et al., 2007).

The process memory retrieval was coded when children explained a particular use generated for a stimulus object by referring to information from semantic or episodic memory, that is, personal past experiences. Across measurement waves, children referred to memory retrieval more frequently than to mental operations when explaining their thinking of novel uses. The comparatively rapid and even accelerated development of the DT process memory retrieval can possibly be attributed to children's development in related developmental domains in this period. First, with growing age and cumulating knowledge and experiences, children may simply have acquired more ideas and associations between ideas to support thinking of novel uses of stimulus objects, that is, their long-term semantic and episodic memories may have expanded and become richer. Attending kindergarten from age 4 introduces children to a new physical, social, and educational context next to family life, which provides new experiences, opportunities for exploration and discovery, and knowledge about the world, enriching children's memory. This explanation is indirectly supported by the finding that memory retrieval was not associated with the originality of mentioned novel uses at age 4 (Bai, Mulder, et al., 2021), while it was at age 6 (Bai, Leseman, et al., 2021), possibly because of a richer semantic and episodic memory. Second, due to progressive language development, children may also have improved in the ability to verbally express their memories and thoughts. The age range from 4 to 6 is marked by rapid development of expressive vocabulary, grammatical knowledge, and pragmatic skills (Hoff, 2009). Indeed, research has shown that children's rapidly growing expressive language skills help them to better express their inner thoughts when they are, as in the present study, supported with cues or interactive dialogues (Flavell et al., 2000). In fact, particularly typical kindergarten settings such as circle time was found to promote children's skill in reflecting upon and talking about their personal experiences (Fivush, 2011; Gathercole, 1998; Nelson & Fivush, 2004). This, in turn, may have helped children to structure their episodic memory and to develop encoding and retrieval strategies that promote retention and recall of richer details of experienced events. Maybe for these reasons, enrolling in kindergarten for some children likely even means a boost of their memory-based DT processes as it was found that children scoring low on memory retrieval at age 4, developed more rapidly than children scoring high at that age (see Figure 4.1).

References to the DT process mental operations occurred much less frequently in children's explanations of their thinking processes compared to memory retrieval across all measurement waves. In addition, mental operations also showed a less steep increase over time than all other measures of DT of the current study. For the development of children's reported use of mental operations, three possible explanations can be considered. First, attending kindergarten also increases children's opportunities for learning about object properties in exploratory play within a well-furnished, stimulus-rich environment, providing children with the perception-action knowledge to manipulate and potentially also simulate disassembling, singling-out, rotating, and recombining object properties in novel ways (Lockman, 2000; Lockman & Kahrs, 2017; Oudgenoeg-Paz et al., 2016). Exploratory play

is a predominant activity setting in Dutch kindergarten (SLO, 2020) and abundant evidence shows a relationship between exploration and spatial cognition (e.g., Newcombe, 2002; Oudgenoeg-Paz et al., 2015). Second, relatedly, performing mental operations on a real or imagined stimulus object and its properties, while resisting the common affordance structure of this object, is likely effortful in nature and requires working memory involvement to hold a mental representation of the object or its properties in active memory in order to mentally operate (e.g., rotate) on this representation, as has been argued in previous research (Bai, Mulder, et al., 2021; Gilhooly et al., 2007; Nusbaum & Silvia, 2011). Therefore, the development of the process of mental operations might also be dependent on children's gradually maturing visuo-spatial working memory (Cowan & Alloway, 2009; Diamond, 2013). Third, like for memory retrieval, children's growing expressive language skills enabled them to reflect and report upon their thinking, which may explain part of the growth in the reported use of mental operations.

Strengths, Limitations, and Future Research

The present study contributes to the current literature on DT in two ways. First, we confirmed the findings of scarce previous studies that DT is already present in children as young as 4 years of age and shows rapid development in the 2 years before children enroll in formal education. Second, to the best of our knowledge, the present study is the first to study the development of DT in young children from a process perspective, focusing on two types of DT processes, memory retrieval and mental operations, identified in previous studies as predicting the fluency and originality of generated uses in the AUT task (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021; Gilhooly et al., 2007). In addition, the robustness and the generalizability of the reported findings were enhanced as we worked with different stimulus sets at consecutive measurement waves (reducing the risk of a testing effect caused by repeated measurements) and used multivariate regression models to analyze theoretically-related outcome measures in the same models (reducing the chance of a type I error) in the present study.

However, the present study also has a number of limitations. One limitation is the psychometric quality of the AUT and the DT process measures in children this young. First, the distributions of some DT measures were skewed, particularly at the earlier measurement occasions and with regard to the DT process mental operations. These results indicate that the AUT itself and the request to reflect upon thinking processes may have been too difficult for children of 4 or 5 years of age, perhaps due to the heavy verbal demands. This was also reflected in the fact that, although being prompted, children did not give a clear and codable explanation for about 50% of the uses generated at each measurement wave. Related to the use of prompting, two other limitations need to be considered. One limitation is that the use of prompts may have introduced subjectivity and experimenter-variance. Although the experimenters were instructed to prompt children to report on their thinking after each generated idea, in practice this was not possible because it tended to break children's thinking

flow and sometimes even upset children. Therefore, the experimenters had to decide in the field when and how many prompts were given in order to keep a balance between prompting children to give explanations and keeping them feel comfortable with the test. Next, specifically the prompts such as “Have you done that before?” may have primed the process memory retrieval and, thus, increased the probability that children would refer to this process to explain how they came up with a particular idea. Given that we applied the same prompting procedure at all four measurement waves, we consider that our conclusions regarding the developmental trend of the DT process memory retrieval still stand. Nonetheless, we propose the following recommendations for future research: (1) A more standardized procedure should be adopted in giving verbal prompts during the testing of the AUT or other DT tasks for obtaining children’s thinking reports. In addition, whether and when to use prompts that may prime some types of thinking processes should be thoroughly thought through, and the effects of such prompts should also be more rigorously examined. (2) In addition to analyzing children’s thinking reports of the AUT to understand the process of DT, a complimentary approach could be analyzing in-depth the exploratory actions performed by children on action-based DT tasks. Studies have shown that action-based DT tasks work well in capturing the full ability range of DT, even in 1- and 2-year-olds (Hoicka et al., 2016, 2018).

The present study was not designed to detect possible fluctuations (e.g., temporary declines or spurts) in children’s DT ability on a more fine-grained time scale, nor to uncover the effects of major transitions, such as the transition from home to kindergarten and from kindergarten to primary school. To gain more insight in the dynamic aspects of DT development and the impact of transitions, it is recommended for future research to apply more frequent measurements and to schedule sufficient measurements before and after major transitions. This type of design could be used to test whether DT development indeed accelerates upon kindergarten enrollment and slows down after the transition to primary school, as has been suggested in past research (Gralewski et al., 2017; Krampen, 2012).

Finally, due to the limited sample size, we could not relate the developmental trajectories of the process measures memory retrieval and mental operations to the developmental trajectories of the product measures fluency and originality. Therefore, it remains uncertain to what extent growth in fluency and originality is possibly caused by growth in the DT processes. In our previous cross-sectional studies with the current sample at age 4 and age 6, we found significant predictive effects of the process measures on the product measures (Bai, Leseman, et al., 2021; Bai, Mulder, et al., 2021). However, future research employing, for example, Latent Change Score modelling (combining growth modelling with cross-lag path analysis) in a larger sample would be needed to establish more firmly whether changes in the DT process measures can be regarded to be causally related to changes in the product measures (for an example of such an approach, see Verhagen et al., 2019).

Conclusion

The present study investigated the development of DT in children from the age of 4 to 6 years using both product and process measures. A linear increase in DT with age was found for the product measures fluency and originality, and for the process measure mental operations. The development of the process measure memory retrieval showed an accelerating increase with age. The present study provides a first glance at the development of the thinking processes involved in DT in young children. Further research is recommended (a) to investigate the development of DT together with the development of other skills that might be involved in DT, such as knowledge, working memory, and language skills, to unravel what drives DT development; (b) to examine DT development around major transitions in children's lives alongside with measures of the environments to which these transitions pertain so as to identify environmental factors that affect the development of DT in young children; and (c) to use more advanced modelling to establish the causal relationships between the development of the DT processes and the development of the DT product measures.

Chapter 5

The Benefits of the Learn to Think Program for Preschoolers' Creativity: An Explorative Study

This chapter is published as: Bai, H., Duan, H., Kroesbergen, E. H., Leseman, P. P. M., & Hu, W. (2019). The benefits of the Learn to Think program for preschoolers' creativity: An explorative study. *Journal of Creative Behavior*, 54(3), 699–711. <https://doi.org/10.1002/jocb.404>

Author Contribution: *Conceptualization: H.B., D.H., and W.H; methodology: H.B., and D.H.; investigation: H.B.; formal analysis: H.B; writing–original draft preparation: H.B.; writing–review and editing: D.H, E.K., P.L., and W.P.; supervision: D.H., E.K., P.L., and W.H.; funding acquisition: D.H. and W.P.*

The Benefits of the Learn to Think Program for Preschoolers' Creativity: An Explorative Study

Abstract

This paper presents the Learn to Think preschool (LTT-P) program for promoting creativity in preschoolers and reviews its potential benefits. LTT-P was designed within the framework of the successful Learn to Think (LTT) creativity program for older students, and both LTT and LTT-P were developed to fit in with the Chinese education system. To assess the potential benefits of LTT-P, a quasi-experimental pretest-posttest control group study was conducted in a preschool in an urban region in the northwest of China, involving 68 middle level and 87 senior level children. The Lines and Circles subtests of the Torrance Test of Creative Thinking were used to examine the development of children's creativity on the dimensions fluency, originality, and elaboration between pretest and posttest. The results suggest that the LTT-P program has the potential to promote young children's creative thinking, especially with regard to the aspects of originality and elaboration. The results for fluency are less clear. To the best of our knowledge, LTT-P is the first program, grounded in a structured learning theory and sound curriculum framework, to support children's creativity development in Chinese early childhood education.

Keywords: Learn to Think, creativity, early childhood

Creativity is considered a key competence in current societies, and education systems across the world are seeking ways to foster creativity in children. According to widely used definitions, creativity refers to the ability to produce original and valuable ideas (Runco & Jaeger, 2012; for ongoing debates of the criteria, see e.g., J.K., Smith & Smith, 2017), which is grounded in the interaction of person, process, and context (Glăveanu, 2013, 2018; Plucker et al., 2004). Early childhood education in daycare centers and preschools is a particularly important context in this regard, for three reasons. First, young children, compared to older children and adults, show still a relatively high level of creativity and related characteristics such as flexibility and curiosity (Glăveanu, 2011), competences which, however, have been found to decrease over time in formal education (Kim, 2011). Second, the predominantly child-centered pedagogy in early, informal compared to later, formal education offers more possibilities to support creativity development (Doddington & Hilton, 2007). Third, early childhood education has been found to generate long-term impact on several skills and competences to the benefit of individuals and society, in particular on “soft” competences such as social skills and behavioral self-regulation (Heckman & Kautz, 2012), and possibly also creativity (e.g., Runco et al., 2010).

The early childhood education system in China is currently struggling with the implementation of creativity education. The traditional emphasis on learning outcomes and the increasing use of standard tests have led to a dominance of direct instruction in (pre)academic skills in early education, often at the expense of activities that could promote creativity (Zhu & Zhang, 2008). In 2012, to change this situation, the Chinese Ministry of Education released the Early Learning and Development Guidelines for 3- to 6-year-old Children, highlighting the need “to respect and protect children’s curiosity and learning interests, helping them to build up learning habits such as taking initiative, being focused, daring to face difficulties, to explore, and to try, and being willing to imagine and create” (p.2). Yet, preschools and especially preschool teachers are observed to have many difficulties with implementing these national guidelines, indicating that they need concrete support (Gao & Huo, 2017)

One way to support preschools is to provide them with well-designed programs for fostering creativity in young children. A number of recent studies from different countries describe and evaluate the results of such programs for fostering creativity in children in the preschool age (e.g., Abdullah Mirzaie et al., 2009; Alfonso-Benlliure et al., 2013; Chronopoulou & Riga, 2012; Dziedziewicz et al., 2013; Garaigordobil & Berrueto, 2011; Gupta, 2008; Smogorzewska, 2012, 2014). The present study adds to this growing body of evidence, introduces a new program for stimulating creativity in preschoolers, the Learn to Think program for Preschoolers (LTT-P), and reports the results of a quasi-experimental evaluation of LTT-P. LTT-P was designed within the framework of the existing Learn to Think (LTT) program for older children. Both LTT and LTT-P were designed to fit in with the context of the Chinese education system, in particular with regard to the usual classroom setting, predominant pedagogy, and common curriculum contents.

Brief Introduction of LTT

LTT was originally designed by Hu et al. (2011) to stimulate students' creativity through the cultivation of their thinking ability. The LTT program includes instruction and practice exercises for different thinking methods across several content domains and is generally conducted throughout the entire school year. A 4-year experimental study on the effectiveness of LTT, involving 166 children from grade one to three, revealed that LTT promoted children's thinking ability and academic achievement. Subsequent studies showed that LTT also facilitated the development of scientific creativity in secondary school (Hu et al., 2013) and other creativity-related competences such as meta-cognitive learning strategies (Hu, Jia, Liu et al., 2016) and learning motivation (Hu, Jia, Plucker et al., 2016). Following the positive results, LTT was integrated in the curriculum of over 300 schools in China (Hu et al., 2013), covering grade one to eight.

Thinking Ability Structure Model (TASM) and the Design of LTT Activities

LTT was built upon the Thinking Ability Structure Model (TASM), proposed by Hu and colleagues (2011). In TASM, thinking ability is regarded as multifaceted, comprising three dimensions: (a) thinking content, i.e., the knowledge that is embedded in different learning materials, and the outcomes of reasoning with this knowledge; (b) thinking methods, i.e., the different kinds or strategies of thinking and reasoning (e.g., divergent, inductive, deductive, imaginary, et cetera); and (c) thinking quality, referring to the degree of fluency, flexibility, critical evaluation, originality, and profundity of the thinking of a person. LTT provides a series of activities across different content domains (thinking content dimension), which stimulate the use of a variety of thinking methods (thinking method dimension) to improve the quality of students' thinking. As described in Hu et al. (2011), the two main dimensions of TASM specify in theory $6 \text{ (thinking content domains)} \times 22 \text{ (thinking methods)} \text{ cells}$, 132 in all, that form the basic structure for designing the activities. However, only the cells deemed most important according to teachers were subsequently transformed into LTT activities per grade level.

Features of LTT Conducive for Creativity

In addition to the activities directly designed to foster creative thinking, a number of general pedagogical principles were included in LTT to ensure successful delivery. These principles can be regarded as conducive for creativity as well (Hu et al., 2013): LTT (a) stresses the use of teaching strategies that introduce cognitive conflict, increase students' motivation and evoke their meta-cognition; (b) emphasizes the importance of creating a social-emotionally safe classroom climate when implementing thinking activities; and (c) encourages students to transfer learned thinking methods to new content domains. These principles are included in a manual for the teachers and addressed in the teacher training program of LTT through lectures, workshops, and teaching simulations. In practice, teachers typically start an LTT activity by introducing the topic, meanwhile stimulating students' interest and motivation; after the introduction, the teachers present examples that contradict

students' prior knowledge to evoke cognitive conflict; subsequently, students are given opportunities to work in groups and to discuss their thinking process with group members; at the end of an activity, teachers ask students to reflect and summarize the thinking methods they have learned in the activity to practice meta-cognitive skills. After each activity, additional enriching content is provided, usually in the form of an extra task, which requires the students to apply the learned thinking method to a different content domain.

LTT-P: The Design and the Adaptations Made to LTT

In view of the effectiveness of LTT for promoting creativity and the wide acceptance of the program by schools, a preschool version of the LTT program was designed, entitled LTT-P. LTT-P adopted TASM as the theoretical foundation, but used a simpler structure based on the two dimensions thinking method and content domain (see Appendix 5.A). Thinking methods that would require a high level of abstraction or cause a high cognitive load, such as philosophical thinking, were excluded to match the developmental level of preschoolers. The contents were mainly selected based on the national guidelines for stimulating 3- to 6-year-old children's development (Chinese Ministry of Education, 2012) and on common subjects in preschool education. Eventually, 20 activities were designed for children at the middle and the senior level in preschool, respectively, 40 in total (see Appendix 5.B).

In addition to TASM as theoretical foundation, LTT-P used the same general pedagogical principles for conducting the activities, with a few adaptations. One adaptation was to put stronger emphasis on independent exploration on beforehand as a way to familiarize children with the learning materials and content domain of a particular activity. For each activity, teachers provided children with ample opportunities to explore the materials and encouraged them to think about the questions or problems posed before starting the group work. Additionally, several procedures were adjusted to reduce the demands on children with regard to abstract discussion and meta-cognition in the activities. First, story-based scenarios were used in almost all activities with concrete materials and drawings, picturing the scene. Second, in addition to evoking cognitive conflicts, modeling by the teacher was used in most activities to help children understand the particular thinking method at stake. For example, in the activity of *Finger Drawing*, the teacher first demonstrated how she used finger prints to paint an animal or an object before asking the children to do this. Third, teachers implementing LTT-P explicitly supported children in their cooperation, showing them how to listen to each other, how to express their own ideas, and how to react to others' ideas. Finally, after an activity was completed, children were encouraged to share what they had learned and how they appreciated the activities as an age-appropriate way to promote meta-cognition. In addition, instead of asking the children to do this as in LTT, the teacher summarized and repeated the targeted thinking method at the end of an activity.

The Present Study

The main purpose of the present study was to assess the potential contribution of LTT-P to the development of creativity in preschool children. The study used a quasi-experimental

design and was conducted in two classes at the middle, respectively the senior grade level in one urban preschool in the northwest of China. LTT-P was administered during one school year to a randomly selected group of children in one class at each grade level and compared to control children in the same and the other class. Being an age-adjusted version of the successful LTT, LTT-P was expected to stimulate the development of creativity in preschool children.

Method

Participants

The participants were recruited from a provincial key preschool located in an urban area of the northwest of China. This preschool is affiliated to and funded by the public university, which hosts the research lab that conducted the current study. Compared to the majority of preschools in China, which are privately owned and usually do not receive public funding, the preschool involved in this study employs better qualified teachers and has access to ample educational facilities and resources. However, the preschool shares typical characteristics with most Chinese preschools, especially regarding structural conditions such as class size and child-to-staff ratio, educational goals, learning contents, and applied pedagogical and instructional strategies. Similar to the mainstream of preschools in China, the participating preschool admits children to three grade levels: 3- to 4-year-olds are admitted to the junior level, 4- to 5-year-olds to the middle level, and 5- to 6-year-olds to the senior level. The LTT-P program was designed for children at the middle and the senior level.

Four classes in total, two at each grade level, were included in this study. At each grade level, one class was chosen to conduct the LTT-P activities. Most LTT-P activities were designed for small group teamwork. It was not feasible to include all children in the class, with a size of around 45, in the LTT-P activities, given the time and energy needed to manage the class. Therefore, children within the classes chosen for the implementation of LTT-P were randomly assigned to either the experimental (participation in LTT-P) or the within-class control condition. By randomly assigning children from the same class to either the experimental or the control condition, class-related effects on the outcome measures could be controlled. To control for possible spillover effects, implying that children of the within-class control condition may increase in creative thinking due to daily interaction with children of the experimental condition, children of two other classes were assigned to the between-class control condition, one class per grade level. In these classes, none of the children received LTT-P.

Eventually, a total number of 182 children participated in the study. Passive parental consent was obtained for virtually all children in all classes. The data of 27 children had to be excluded for several reasons: (a) 25 children did not have complete data (either pretest or posttest data were missing); (b) one child in the experimental group at the senior level participated in only six out of 20 LTT-P activities, which was considered insufficient; (c) one child in the within-class control group at the middle level was much older (67 months) than

the average and was reported by the teacher to be a child with special educational needs. The final sample included 155 children (92 boys). Table 5.1 includes information with regard to the age and gender of the children, and the final sample size by grade level and condition.

Measures

The Lines and Circles subtests of the figural Torrance Test of Creative Thinking (Torrance et al., 1992/2008) were used at, respectively, pre- and posttest to evaluate children's creativity. The forms used at pre- and posttest differed but are psychometrically equivalent and can be considered parallel forms (Torrance et al., 1992/2008). TTCT is one of the most widely used tests for assessing creativity. Test-retest reliability of TTCT is reported to exceed 0.80 (Torrance, 1972), and the test scores correlate highly with individuals' creative behaviors and personality traits as measured by an established scale (Garaigordobil & Berruero, 2011).

The test at pretest presented children with pairs of lines, and the test at posttest presented children with circles. For both tests, children were requested to make as many drawings as possible while integrating the presented stimuli in their drawings (Torrance et al., 1992/2008). The tests took 10–20 minutes to finish. No strict time limit was set in this study, given the age of the children, allowing them sufficient time to express their thinking in the drawings. Testing was stopped when children repeated previous drawings several times or when children indicated that they wanted to stop.

Trained research assistants administered the tests to groups of four to five children. The assistants were intensively instructed, performed trial tests on each other, and were closely supervised by the first author during the first test session. At pretest, due to practical limitations, no separate test room was available at the preschool, and thus, the tests were administered in the classroom during one session. At posttest, the tests were administered in a separate room during several short sessions. Before starting the actual test, one training item was presented, and the research assistants demonstrated how to make a drawing (a pair of lines was extended to a pencil and a circle was extended to an apple) in order to ensure that children understood the test. When the children had completed the test, they were asked to name all their drawings. If the name could not unequivocally be related to the drawing, children were asked to explain the drawing in detail. These names and explanations were recorded by the research assistants for later scoring.

Children's drawings were scored on the common aspects fluency, originality, and elaboration (Torrance et al., 1992/2008). Fluency was scored following the scoring manual (Torrance et al., 1992/2008). Repeated drawings and drawings that clearly did not use the given stimulus were excluded from the scoring. For each child, a fluency score was calculated as the total number of the remaining drawings. Originality was scored based on the frequency of the ideas expressed by a child in his or her drawings (as also indicated by the names or the more detailed explanation given) relative to all ideas expressed by all children in the whole sample: the more infrequent relative to the whole sample, the more original. Drawings that appeared in the tests of less than 10% of the children were scored 1 on originality, all other

drawings 0. We used a relative measure of originality rather than the established norm of the test, because this norm is based on a US sample and might possibly not fit the Chinese background of the children. For drawings in which two or more of the given stimuli were combined, bonus scores for originality were given, in accordance with the scoring manual. Eventually, an overall originality score per child was computed as the sum of the originality scores of all valid drawings plus the bonus scores. Finally, elaboration was scored based on the overall impression of the richness of details in a child's drawings rated on a 5-point scale (1 = *not elaborated at all*; 5 = *very elaborated*). Scoring was done by the first author.

Limitations of the Design

It should be noted on beforehand that the current design suffered from several limitations: (1) There was only one preschool involved, which moreover benefitted from its relation with the public university. The findings of this study, therefore, cannot be easily generalized to the entire system of preschool education in China. (2) The experimental group and the within-class control group were in the same class throughout the whole school year, except for the LTT-P activities. This could lead to a spillover effect, possibly obscuring the experimental effect. (3) Only one researcher scored children's creativity products. Although several procedural measures were taken to assure the objectivity and reliability of the scoring, coder bias cannot be ruled out.

Procedures

The study was conducted in the semesters before and after the winter break, spanning a total period of around 9 months. The winter break lasted 1 month. The program provided activities for around 6 months. Depending on the preschool schedule, one or two activities (of 30-45 min) per week were conducted at each grade level. The first author conducted the program under the supervision of the preschool vice principal of teaching and the second author. All children took the pre- and posttests, but only children of the experimental group took part in the LTT program activities. Pretest and posttest were conducted two weeks before, respectively two weeks after the LTT-P program,.

Results

Means and standard deviations of fluency, originality, and elaboration by condition are presented in Table 5.1, separately for the pre- and posttest. Repeated measures multivariate analysis of variance (MANOVA) was applied to analyze the data of the middle and the senior level, respectively, to assess differences between the experimental condition and the two control conditions in the increases of the creativity measures fluency, originality, and elaboration between pre- and posttest. For both grade levels, the Box's tests were significant ($p_{\text{middle}} = .017$ and $p_{\text{senior}} = .011$), indicating that the assumption of equality of covariances across groups was violated. Furthermore, Levene's tests were not significant for the creativity measures at pretest, but significant for all measures at posttest. Altogether, these results suggest that implementation of LTT-P enlarged the differences between children in the

measures between the conditions. Wilks' Lambda's were used to interpret the results, because the Box's tests were not significant at $p < .001$ and the sample sizes were fairly equal across groups (Mertler & Vannatta, 2013). The results of MANOVAs and the post hoc ANOVAs are presented in Table 5.2.

The results of the MANOVA for the middle level, Wilk's $\Lambda = .78$, $F(6, 126) = 2.86$, $p < .05$, $\eta_p^2 = .12$, showed a significant multivariate Condition \times Time interaction effect on fluency, originality, and elaboration. Subsequently, post hoc univariate ANOVAs revealed condition effects on the gains of fluency, $F(2, 65) = 3.83$, $p < .05$, $\eta_p^2 = .11$, originality, $F(2, 65) = 3.32$, $p < .05$, $\eta_p^2 = .13$, and elaboration, $F(2, 65) = 5.15$, $p < .01$, $\eta_p^2 = .14$. Finally, pairwise comparisons were conducted with a Bonferroni corrected significance level of $p < .05/3 = .018$. The comparison results showed that (1) children in the experimental group had a significantly larger increase in fluency, $F(1, 35) = 6.16$, $p = .018$, but not in originality and elaboration, than children in the within-class control group; (2) children in the experimental group had significant larger increases in originality, $F(1, 50) = 6.02$, $p = .018$ and elaboration, $F(1, 50) = 9.09$, $p = .004$, but not in fluency, than children in the between-class control group; and (3) children in control groups did not differ in any measure.

The results of the MANOVA for the senior level, Wilk's $\Lambda = .74$, $F(6, 164) = 4.54$, $p < .001$, $\eta_p^2 = .14$, showed also a significant Condition \times Time interaction effect. The post hoc ANOVAs revealed condition effects on the gains of fluency, $F(2, 84) = 3.11$, $p < .05$, $\eta_p^2 = .069$, originality, $F(2, 84) = 6.16$, $p < .01$, $\eta_p^2 = .13$, and elaboration, $F(2, 84) = 6.55$, $p < .01$, $\eta_p^2 = .14$. Pairwise comparisons showed that (1) there was no significant difference between children in the experimental group and the within-class control group; however, (2) children in the experimental group had significant larger increases in originality, $F(1, 63) = 10.81$, $p = .002$, and elaboration, $F(1, 63) = 14.13$, $p < .001$, but not in fluency, than children in the between-class control group, and (3) children in the within-class control group had a significant larger increase in originality, $F(1, 61) = 6.32$, $p = .015$, than children in the between-class control group.

In summary, children at the middle level who participated in LTT showed larger increases in fluency than children in the within-class control group, and in originality and elaboration than children in the between-class control group. Children at the senior level showed larger increases in originality and elaboration than children in the between-class control group, but not than children in the within-class control group. Regarding the two control groups, no differences were found in increases in creativity at the middle level. At the senior level, children in the within-class condition showed a larger increase in originality than children in the between-class condition.

Chapter 5

Table 5.1

Sample Descriptives and Means and Standard Deviation of Creativity Measures, N = 155 (92 boys)

Groups	EXP				Ctrl-S				Ctrl-NS			
Middle level												
<i>N</i> cases	21 ^a				16				31			
<i>N</i> boys (%)	11 (52.38%)				8 (50.00%)				18 (58.06%)			
Mean age in months (<i>SD</i>)	54.10 (2.92)*				54.25 (2.84)				57.35 (3.77)			
	Pretest		Posttest		Pretest		Posttest		Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fluency	4.52	4.16	14.67	8.26	4.41	4.65	9.71	4.22	5.39	3.58	11.94	4.65
Originality	3.67	4.54	8.76	6.84	3.18	3.68	5.65	4.89	4.1	3.94	5.03	3.29
Elaboration	1.33	.73	3.00	.89	1.35	.79	2.59	.80	1.74	.86	2.71	.78
Senior level												
<i>N</i> cases,	24				22 ^b				41			
<i>N</i> boys (%)	15 (62.50%)				14 (63.63%)				26 (63.41%)			
Mean age in months (<i>SD</i>)	68.00 (3.40)				66.84 (3.45)*				69.27 (3.74)			
	Pretest		Posttest		Pretest		Posttest		Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fluency	10.38	7.76	19.75	8.36	8.41	5.54	17.77	10.70	13.8	6.02	18.71	3.84
Originality	8.92	9.56	17.58	10.48	6.55	5.82	13.18	13.10	11.8	8.47	11.15	5.80
Elaboration	2.21	.78	3.67	.70	2.09	0.97	3.05	.38	2.66	.79	3.32	.61

Note. EXP = Experimental group; Ctrl-S = Within-class control group; Ctrl-NS = Between-class control group.

^a One child's age information was missing. ^b Three children's age information were missing. *The mean age and SD were calculated based on available age information of the children.

Table 5.2

Results of the Repeated Measures Multivariate Analysis of Variance (MANOVA) and the Subsequent Analysis of Variance (ANOVA)

Statistics		Middle level	Senior level
MANOVA	Time	Wilk's $\Lambda = .21$, $F(3, 63) = 78.42$, $\eta_p^2 = .79$	Wilk's $\Lambda = .31$, $F(3, 82) = 59.58^{***}$, $\eta_p^2 = .69$
	Group	Wilk's $\Lambda = .90$, $F(6, 126) = 1.15^a$, $\eta_p^2 = .05$	Wilk's $\Lambda = .76$, $F(6, 164) = 4.02^{***}$, $\eta_p^2 = .13$
	Time \times Group	Wilk's $\Lambda = .78$, $F(6, 126) = 2.86^*$, $\eta_p^2 = .12$	Wilk's $\Lambda = .74$, $F(6, 164) = 4.54^{***}$, $\eta_p^2 = .14$
ANOVA	Fluency	$F(2, 65) = 3.83^*$, $\eta_p^2 = .11$	$F(2, 84) = 3.11^*$, $\eta_p^2 = .069$
(Time \times Group)	Originality	$F(2, 65) = 3.32^*$, $\eta_p^2 = .13$	$F(2, 84) = 6.16^{**}$, $\eta_p^2 = .13$
	Elaboration	$F(2, 65) = 5.15^{**}$, $\eta_p^2 = .14$	$F(2, 84) = 6.55^{**}$, $\eta_p^2 = .14$

^a Not significant. * $p < .05$. ** $p < .01$. *** $p < .001$

Discussion

The results of the current study suggest that the development of children's creativity, at least regarding the originality and elaboration of their responses to the widely used Torrance Test of Creative Thinking, can be supported by implementing a program of activities, such as LTT-P, in preschool classrooms that stimulate different ways of thinking across different content domains. LTT-P is a new creativity program for preschoolers at the middle (4- to 5-year-olds) and the senior level (5- to 6-year-olds) in China and was based on the previously developed LTT framework for older children. The current findings add to the evidence on the usefulness of the LTT framework for promoting creativity (Hu et al., 2013) and creativity-related capacities (Hu, Jia, Plucker, et al., 2016; Hu, Jia, Liu, et al., 2016).

As a general trend, the data showed that children who participated in LTT-P gained more in creativity on most outcome measures than children in the between-class control condition. However, with regard to the within-class control condition the results were less clear (at the middle level: no difference regarding originality and elaboration gains; at the senior level: no differences in creativity gains at all). The results for one dimension of creativity, fluency, were also much less clear. We will elaborate on the main findings and offer explanations for the unexpected results below.

The Benefits of LTT-P on Originality and Elaboration

Regarding originality and elaboration, significant differences in growth were found between the experimental and the between-class control groups. These two groups of children were not strictly equivalent as we were not able to work with random assignment, but they can be considered as comparable. The children were from the same social backgrounds and shared a similar educational context (for example, regarding class size, curriculum, educational climate, teacher qualifications). However, we found no significant differences in growth between the experimental and within-class control children. Note that due to within-classroom randomization both groups could be considered as fully equivalent with only the experimental treatment differing between the groups, while other potentially relevant classroom factors were fully shared and thus controlled for. Therefore, the lack of significant differences in gains in originality and elaboration between experimental and within-class control group may point to spillover effects that counter-acted the experimental effect. Although children in the within-class control condition did not participate in the LTT-P program activities, they were in the same classrooms as children who did. The increased motivation and enthusiasm of the experimental children may have influenced the classroom climate and thereby stimulated the creativity of the within-class control children also. Likewise, the training of the experimental children in reflecting and elaborating on ideas and activities may have had a wider influence on classroom processes, impacting also within-class control children. Second, in actual practice, we observed that within-class control children were very curious about what the experimental children were experiencing during

program time and they were found to be very motivated to do the posttests. These effects, of course, were not present in the between-class control group.

The positive influence of LTT-P on originality found in the current study may actually point to a moderating effect of the context of Chinese preschool education, as previous studies in other countries have revealed that creativity programs are not always successful in promoting originality (e.g., Alfonso-Benlliure et al., 2013; Smogorzewska, 2014). Originality, unlike fluency, seems more difficult to stimulate in education contexts in other countries. Chinese education has been thought to hinder students' creativity more than education programs in other countries (Niu & Sternberg, 2001, 2003; Niu et al., 2007), but may for this reason leave more room for creativity promoting programs to yield effect. Several researchers have pointed to Confucianism as a factor in this regard (e.g., Niu & Sternberg, 2001). As a system of values and norms, Confucianism favors the creation and maintenance of a hierarchical social system and emphasizes socialization of proper behavior according to prevailing social norms (Wu & Albanese, 2010). It especially discourages arguments and discussions (Wu, 2004). These ideological features have profoundly influenced the value system of education as well, already in preschool, giving rise to a pedagogical climate in classrooms that prevents students from generating and, in particular, expressing novel ideas. Originality, thus, is hindered or under-stimulated in this context, insofar as originality requires students to cross boundaries, think outside of the box, and express novel ideas. Therefore, an intervention program that explicitly (and intensively) encourages children to generate, express, and share novel ideas may show a relatively large effect. Evidence pertaining to older children indeed indicates that, when given explicit instructions, Chinese students perform better in creativity tasks than their counterparts in other countries (Wong & Niu, 2013).

As for elaboration, the results were largely consistent with our expectations. Children developed spontaneously in elaboration skills, and the creativity program LTT-P added to that significantly, at least when comparing the experimental groups with the between-classroom control groups where no spillover effect could have disturbed the findings. The positive influence of LTT-P on children's elaborated thinking may be due to two pedagogical processes introduced by LTT-P. First, as a key principle, children were explicitly encouraged to describe their thinking process and to explain their ideas to both peers and teachers in almost all activities. Second, collaborative teamwork was an important element of many LTT-P activities, stimulating children to explain and compare their ideas with those of others and to reflect explicitly on commonalities and differences. As a consequence of these two processes, children were trained to be elaborate in both thinking and expressing their ideas.

The Lack of Benefits of LTT-P on Fluency

There were no stronger gains in fluency for the experimental than the between-class control group, but at the middle grade level there was a significantly stronger gain for the experimental than the within-class control group, contradicting the idea of a spillover effect

of LTT-P within classrooms. This finding can be coincidental, but perhaps, this also indicate that the spillover effect, at least for the younger children, was not as strongly affecting fluency as it may have been affecting originality and elaboration. A possible explanation could be that fluency increased rather strongly between the pre- and posttest measurements regardless of condition, and more so at the middle level than at the senior level. This strong spontaneous increase in fluency in all conditions may have attenuated the effect of LTT-P. Note in addition that in the experimental condition the standard deviation at posttest was substantially larger than in the other conditions. This could suggest that the fluency training part of the LTT-P program worked well for part of the children only but less well for the rest, which could explain that an overall noticeable spillover effect on the whole classroom was absent, while the mean scores still differed between experimental and within-class control conditions. Upon closer scrutiny of the data, we indeed found that about half of the children in the experimental group showed substantial gains in fluency, while the other half showed much smaller gains.

The lack of an effect of LTT-P on fluency seem to contradict findings reported in the literature, which suggest that fluency can be easily promoted, as almost all preschool creativity programs were reported to have effects on fluency (Abdullah Mirzaie et al., 2009; Chronopoulou & Riga, 2012; Dziedziewicz et al., 2013; Garaigordobil & Berruenco, 2011) but not always on the other dimensions of creativity (Alfonso-Benlliure et al., 2013; Smogorzewska, 2014). Although the substantial gain in fluency that children especially at the middle level showed in this study, regardless of condition, is in line with the idea that fluency can be relatively easily promoted (and even can increase spontaneously), we did not find a clear condition effect. Possibly, the fluency training component of LTT-P could not add much to the already substantial spontaneous growth in fluency, explaining the difference in findings with other studies.

Conclusion

The present study reviewed the benefits of the Learn to Think preschool program on creativity in children using a quasi-experimental design. LTT-P was designed specifically to enhance the support of creativity and thinking skills in Chinese preschool education. The results suggest that LTT-P has the potential to promote young children's creative thinking, especially with regard to the dimensions originality and elaboration, as was expected based on the results of the LTT program for older children. In the context of the Confucianist pedagogical climate of Chinese preschools, the positive influence of LTT-P on children's originality was regarded as promising, as children in such a climate may not develop the skill to generate and express original ideas spontaneously. However, it should be noted that the results are not fully conclusive as the study design suffered from several limitations.

In addition to the design limitations mentioned in the Method section, the current study showed two other shortcomings that should be improved in future research. First, with regard to measuring creativity, the TTCT, even though widely used, may not be the best choice. As

indicated in the present study and also in the literature, creative thinking as currently defined and operationalized reflects a specific cultural construct. TTCT was built predominantly on the basis of a creativity construct defined within the American culture, which may not do full justice to creativity constructs in other cultural systems. Additionally, in LTT-P much emphasis was put on promoting the processes of children's creativity, whereas the TTCT only measures the products. For future research, it is recommended to use a research strategy that can better capture the thinking processes that lead to creative products and can reveal how contextual factors may influence these processes (for further suggestions see John, 2017 and Van Dijk et al., 2019). Such a research strategy could entail qualitative analysis of children's explanations of how they came up with particular solutions.

A second shortcoming of the current study is the use of a cross-sectional design to examine the added value of LTT-P for children at the middle and the senior level separately. LTT-P, however, is deliberately designed as one comprehensive program for both levels with a hierarchical goal structure specifying different age-appropriate but over time coherent goals for younger and older children. In the current study, however, we could not examine whether the hierarchical structure was indeed appropriate and coherently spanned both grade levels. Future research should take this into account and preferably use a longitudinal design with a larger time span.

Notwithstanding these limitations, we argue that this study is valuable as it provides the first indications that a structured program to promote creativity in preschool children can be effective, at least in the context of the Chinese early education system. To the best of our knowledge, LTT-P is the first program, grounded in a structured learning theory (i.e., TASM) and a sound curriculum framework (i.e., LTT), to support the creativity development of preschoolers in China. Implementing LTT-P is likely beneficial for promoting originality in children's thinking, and this makes the program of particular interest for preschools that attempt to reform their education programs. In addition, implementing LTT-P may also contribute to the prevention of early gaps (educational inequality in China, e.g., Luo & Li, 2017) in the development of soft skills like creativity between children attending common preschools with a traditional pedagogy and children attending elite private preschools or public funded preschools, where open, child-following and soft skills supporting pedagogies have already been introduced (e.g., Wang & Chen, 2014). The accessibility, transferability and low costs of LTT-P, compared to creativity programs imported from other countries, can contribute to more equity in preschool education in China.

Chapter 6

General Discussion

Divergent thinking (DT) can be defined as a way of thinking that enables individuals to explore or to think in multiple directions (Guilford, 1956; Wang et al., 2017). As an important component of creativity, DT is considered to be particularly relevant for generating original or novel ideas (Runco & Acar, 2012) and, thus, has received considerable research interest. However, available DT research focuses mainly on adults and occasionally on school-age children. There is a scarcity of research that specifically looked into DT in young children, even though recent studies have shown that DT emerges in children as early as 1 or 2 years of age (Hoicka et al., 2016, 2018). The primary aim of this thesis, therefore, was to provide answers to questions such as how young children come up with original ideas during DT and how DT develops in early childhood, which are currently understudied. To this end, a cohort of 107 children was followed for 1.5 years, from the age of 4 to 6 years. At four measurement waves, children were measured on DT with the Alternative Uses Task (AUT; Gilhooly et al., 2007; Guilford, 1967); and at the last measurement wave these children were also administered a battery of standard executive function (EF) tests. To study how children come up with original ideas during DT, we first investigated the process of DT in young children by analyzing children's thinking reports, which were obtained while children were performing the AUT (Chapter 2). The cohort data on DT obtained at the first measurement wave were used in this study. The results indicated that top-down, effortful executive processes played a role in the generation of original ideas during DT. Therefore, a subsequent study examined specifically how top-down, effortful executive processes operate during DT, using the cohort data on both DT and EFs obtained at the last measurement wave (Chapter 3). To study how DT develops in children, the developmental trajectories of the DT measures as derived from the AUT were modeled using the longitudinal data (Chapter 4). Next, apart from answering the aforementioned two fundamental questions, this thesis aimed to contribute to applied research regarding fostering creativity in children and, in particular, to take up the challenge the Chinese education system is currently facing in the implementation of creativity education in early childhood. For this purpose, a creativity intervention program, i.e., the Learn to Think program for Preschoolers (LTT-P), was designed and evaluated with a quasi-experiment with pre- and posttests in the context of Chinese early childhood education (Chapter 5).

In this final chapter, we first summarize the main findings of studies reported in this dissertation. Subsequently, we discuss the theoretical and practical implications of the findings and give suggestions for future research. Finally, we address the limitations of this dissertation and put forward the main conclusions.

Summary of the Results

How do Children Come up With Original Ideas?

While performing a DT task, people often begin with generating many mundane and conventional ideas in a short time frame and only gradually shift to generating more original ideas but at a slower pace. This phenomenon is named the serial order effect of DT and has

Chapter 6

been reported in a range of samples, varying from children as young as 7 years old to adults (Christensen et al., 1957; Mednick, 1962; Milgram & Rabkin, 1980; Wang et al., 2017; Ward, 1969). Chapter 2 and Chapter 3 reported studies that analyzed the process of DT in children and addressed the roles of domain-general executive processes and domain-specific DT processes in the serial order effect, in order to deepen our understanding of what creativity in general and DT as a key component of creativity are about.

Chapter 2 reported on 4-year-olds' thinking processes underlying the generation of uses during the AUT. Within the AUT, children were instructed to give as many unusual uses as possible for common objects and to report on their thinking processes in relation to the generated uses. Children's utterances were analyzed using a specially designed coding scheme. This analysis revealed that children were not always able to explain how they arrived at a particular use. Explanations were missing or not codable for 57% of the uses mentioned by children. When children did explain how they came up with an idea, their explanations were mostly indicative of thinking processes involving retrieval of semantic or episodic long-term memory. These type of processes, referred to as "memory retrieval" from here on, are thought to be associative in nature and occurred predominantly automatically and in a bottom-up manner (Beaty & Silvia, 2012; Mednick, 1962; Gilhooly et al., 2007; Sowden et al., 2015). Occasionally, however, children's explanations were also indicative of other types of thinking processes that are more effortful, top-down, and executive in nature. These processes included *Performing mental operations on a stimulus* (e.g., disassembling, re-assembling, turning, distorting, folding; referred to as "mental operations" from here on), *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; see Chapter 2, Table 2.1, for definitions and examples of the DT processes). In addition, it was revealed that both associative and executive DT processes positively predicted the fluency of DT (i.e., the average number of uses generated for a stimulus object), while only the executive DT process mental operations uniquely and positively predicted the originality of DT (i.e., the extent to which the generated ideas are statistically rare or rated as original).

This study also investigated how the associative and the executive processes relate to DT while taking into account how the thinking processes unfolded in real time. To this end, children's generated uses were treated as a quasi-time-stamped variable based on the rank number of each use in the sequence of all generated uses per stimulus object. It was found that children of 4 years of age already showed a serial order effect in the originality of their responses, that is, the originality of uses increased the later a particular use was generated in the series of mentioned uses per object. In addition, similar serial order effects characterized the occurrence of all executive DT processes, showing that the later in the sequence a

particular use was generated, the more likely it was that children would explain how they arrived at this use by referring to these processes. In particular, the graphs representing the serial order effects of originality and the process of mental operations revealed similar timing patterns. Regarding the associative process of memory retrieval, a reversed but non-significant serial order effect was present: the probability that children referred to this type of process was initially high, but after the first few mentioned uses, this probability declined with increasing use rank within a stimulus object. These findings suggested that increasing originality particularly depends on increasing involvement of the executive process of mental operations during DT, which involves mentally isolating properties or parts of objects and the subsequent recombination of these parts and properties into a new structured whole.

Chapter 3 replicated and expanded on the study reported in Chapter 2 by investigating the role of executive processes during the real-time unfolding of original ideas within the AUT. The same sample of children but now at the age of 5 to 6 years (1.5 years later) were assessed on DT using the AUT in the same way as in the study reported in Chapter 2. Each generated use was coded on originality and on whether children's explanations of how they came up with this use referred to the two most important DT processes identified in the previous study: *memory retrieval* and *mental operations*. It was found that there was a serial order effect in the originality of uses and that mental operations positively predicted the originality of uses, which replicated findings of Chapter 2. However, apart from mental operations, this study also found a significant positive effect of memory retrieval on the originality of uses, which was absent in the previous study. Similarly, such a positive effect of memory retrieval on originality was also reported in a study among university students by Gilhooly et al. (2007). We inferred that, after successive testing with the AUT and due to being older and getting more used to the structured context of kindergarten, children may have understood the task demands better and complied more with the instruction of coming up with real novel uses, thus thinking in a manner more similar to adults. As a consequence, children at age 5 to 6 years (and adults alike) may have approached the AUT differently compared to when they were 4 years of age, showing likely a developmental change. Additional support for this interpretation came from the findings regarding the timing of DT processes involved in the generation of original ideas. In Chapter 2, we proposed that 4-year-olds may gradually shift from less original responses based on automatic associative memory processes to comparatively original responses based on controlled mental operations during idea generation for a stimulus object. However, in Chapter 3 both DT processes memory retrieval and mental operations *negatively* interacted with use rank in predicting the originality of uses (this effect was only borderline significant for memory retrieval), showing that the effects of memory retrieval and mental operations on originality were already present in the initial stage, remaining stable for a while, but declining thereafter and turning even into a negative effect for the last 3% of all generated uses. Therefore, children of 5 to 6 years of age who better understood the task requirement of generating something *new* and *original* may have shifted almost immediately to the use of controlled executive processes like mental

Chapter 6

operations, and maybe also more effortful forms of memory retrieval (e.g., recalling detailed, unique personal memories), instead of initially relying on spontaneous memory-based associations as they did when they were 4-year-olds.

In addition to analyzing how the domain-specific DT processes memory retrieval and mental operations operate during DT, the study reported on in Chapter 3 also investigated how domain-general EFs relate to and interact with DT processes during the unfolding of original ideas within the AUT. To this end, children were assessed on the domain-general EFs inhibition, shifting, working memory, and selective attention, using standard tasks. We had expected that higher EFs would facilitate the use of executive processes and consequently the generation of original ideas over time, based on the idea that increasing originality depends on increasing involvement of effortful executive processes during idea generation for an object, as put forward in Chapter 2. However, contrary to our expectations, none of the EFs were predictive of originality, nor did any of these EFs interact with either memory retrieval or mental operations in the prediction of the originality of uses. In addition, selective attention (but none of the other EFs) was even found to negatively interact with use rank in the prediction of originality, showing that *low* selective attention was generally beneficial for the generation of original ideas, and children with lower selective attention produced substantially more original responses in a later stage of idea generation for a stimulus object than children with higher selective attention. Possibly, lower selective attention facilitates perceiving a broader set of external stimuli (e.g., irrelevant objects in the test environment; cf. Van Dijk et al., 2020) or promotes a broader spreading of activation to more remote and unrelated memory representations in semantic and episodic memory (Benedek, Könen, and Neubauer 2012; Brown 1973; Mednick 1962), thereby leading to the generation of more unique, and thus original, ideas.

Taken all findings of Chapter 3 together, there seemed to be a complex interplay of different cognitive processes during DT in children at age 5 to 6 years: controlled effortful processes early in the sequence of generating novel uses may be associated with children's compliance with the task demand of generating something new and novel; while a broader attentional scope (i.e., a relative lack of focus), may gradually replace the controlled processes, possibly due to depletion of the latter, and facilitate idea generation towards the end.

How Does DT Develop in Children?

Chapter 4 reported on young children's development of DT from 4 to 6 years of age. One hundred and seven children were assessed with the AUT task approximately every 6 months, four times in total, spanning a period of 1.5 years from the time children entered kindergarten until shortly before they enrolled in formal primary education. The AUT was administered as described above. Children were instructed to generate as many unusual uses of six common objects while they were also asked to report on their thinking processes. Based on the ideas children generated, again two product measures were computed at the child level:

fluency and originality. Also in this study, the DT processes memory retrieval and mental operations were coded from children's explanations. Based on the coding, two process measures of memory retrieval and mental operations were computed at child level and were used for the analysis. The developmental trajectory of DT was investigated through the modelling of the linear and quadratic effects of age on both the product and the process measures. The results revealed substantial growth in fluency and originality with age. Maybe due to the active prompting as applied in the testing that encouraged children to spend more effort, we found that the average fluency score of children of 5- to 6-year-old, measured at the last measurement wave, was comparable to what has been reported in school-age children (Van Dijk et al., 2020) and adults (Gilhooly et al., 2007). In addition, increases with age were also observed in the number of references children made to the DT processes memory retrieval and mental operations when explaining their thinking.

How Can We Foster Creative Thinking?

Following the in-depth investigations of the process of DT in young children reported on in Chapters 2–4, Chapter 5 presented a more applied study on stimulating creativity in young children. In this chapter, a structured intervention program, i.e., the Learn to Think program for Preschoolers (LTT-P), was introduced, and the results of a field study that reviewed the potential benefits of LTT-P in facilitating creativity in children of 4 to 6 years of age were reported. LTT-P was designed within the framework of the LTT creativity program, which was proven to be effective in facilitating primary and secondary school students' creative thinking (Hu et al., 2011) and other creativity-related competences such as meta-cognitive learning strategies (Hu, Jia, Liu, et al., 2016) and learning motivation (Hu, Jia, Plucker, et al, 2016). Previous research has suggested that Chinese education is more likely to hamper creative thinking in students than education programs in other countries (Niu & Sternberg, 2001, 2003; Niu et al., 2007). Therefore, both LTT and LTT-P were specifically designed to fit in with the Chinese education system in order to support the implementation of creativity education therein, taking context-related factors including the usual classroom setting, the mainstream pedagogical and didactical approaches, and the common curriculum contents into account. In China, early childhood education is predominantly provided by preschools at three grade levels: 3- to 4-year-olds attend the junior level, 4- to 5-year-olds attend the middle level, and 5- to 6-year-olds attend the senior level. LTT-P consists of activities for children at the middle and senior levels. A field study using a quasi-experimental design with pre- and posttests was carried out to assess the benefits of LTT-P, involving 68 children from the middle level and 87 children from the senior level. The quasi-experiment spanned two semesters, and one or two activities were conducted every week at each grade level for about 6 months. The Line and Circle subtests of the figural Torrance Test of Creative Thinking, which are psychometrically equivalent and thus considered as parallel forms (Torrance et al., 1992/2008), were used respectively in the pretest and the posttest to assess the fluency, originality, and elaboration of children's creative thinking. The results of repeated measure multivariate analysis of variance showed similar trends for children from the two

grade levels, suggesting that the LTT-P has the potential to promote young children's creativity in the dimensions of originality and elaboration; nonetheless, the results on fluency were less clear-cut, leaving it unclear whether LTT-P is also beneficial for promoting the creativity dimension of fluency.

Theoretical Implications

Understanding DT: A Complex Interplay of Different Types of Cognitive Processes

The serial order effect of DT is regarded as a robust paradigm for studying the process of DT and in particular the emergence of original ideas therein (Beaty & Silvia, 2012; Gilhooly et al., 2007; Hass, 2017b; Wang et al., 2017). The findings of the studies reported in Chapters 2 and 3 confirmed that the serial order effect is present in children as young as 4 to 6 years of age. In addition, the in-depth analysis of children's DT processes also showed remarkable similarities with what has been reported in a study with adults (Gilhooly et al., 2007).

An important implication of our findings relates to two explanations regarding the occurrence of the serial order effect of DT. First, the *associative* account, grounded in the early work of Mednick (1962), explains the serial order effect by referring to the spreading activation in semantic and episodic long-term memory (Collins & Loftus, 1975; Conway & Engle, 1996). As for the AUT, for instance, when subjects are given a stimulus object and are asked to generate as many original uses as possible, their semantic memory in relation to this object is automatically activated, driven by associative, bottom-up, and relatively effortless processes, according to this account. The activation first pertains to well-entrenched near connections of the stimulus and, thus, likely leads to more conventional and familiar uses. This process gradually spreads to less well-entrenched, "weaker", and more distal connections of the stimulus that are more likely to result in unconventional and novel ideas. In contrast to the associative account, the *executive* account, emerging from more recent studies, suggests that the occurrence of the serial order effect is the result of the increasing involvement of executive, top-down, and effortful processes in DT over time (Beaty & Silvia, 2012; Hass, 2017a, 2017b; Silvia, 2015). Following this line of reasoning, executive processes may be conducive to the generation of original ideas by (a) facilitating the activation and the effective retrieval of semantic elements that are only remotely associated with the stimulus (Benedek & Neubauer, 2013) and/or (b) enabling thinking processes that are no longer based on finding (remote) memory associations but on manipulating mental representations and creating new links between (parts of) representations (Gilhooly et al., 2007).

The findings reported in Chapters 2 and 3, although both confirming the simultaneous involvement of associative and executive processes in DT, showed differences with regard to the timing and the effects of these processes during DT. The findings reported in Chapter 2 were in line with the theoretical model summarized above and suggested that DT in children of 4 years of age initially involved generating ideas based on automatic associative retrieval

from long-term memory, while gradually, when memory-based uses were becoming exhausted, the controlled, effortful thinking processes were called in. More specifically, the increasing originality of uses over the course of idea generation on a stimulus object likely depended on the increasing involvement of effortful executive processes, especially the process of mental operations. However, the findings reported in Chapter 3, based on the same sample but now at age 5 to 6 years, suggested that children of this age approached the AUT differently than when they were 4 years of age. This was evident in two ways. First, instead of initially relying on spontaneous retrieval of long-term memory as most of them did so at age 4 years, children at age 5 to 6 more likely shifted almost immediately to controlled executive processes. This interpretation was underscored by the findings that both memory retrieval and mental operations positively predicted the originality of uses and that these effects were already present in the initial stage of idea generation and remained stable for the generation of the majority of uses. The significant positive prediction of originality by the process memory retrieval, found at age 5 to 6, also implied that memory retrieval may not solely rely on automatic associations of well entrenched, nearby elements in the semantic and episodic memory networks, as we had initially assumed (see Chapter 2); instead, memory retrieval may also involve effortful thinking, possibly through inhibiting the less original responses emerging from mundane memories. Second, apart from the executive processes, children likely also benefitted from a broader attentional scope in the generation of original ideas, particularly at a relatively late stage of idea generation, supported by the finding that selective attention negatively interacted with use rank in the prediction of the originality of uses. Taking together, the somewhat conflicting findings of Chapters 2 and 3 led to the tentative conclusion that neither the associative nor the executive account alone can fully explain the serial order effect of DT. Instead, there might be a complex interplay of different cognitive processes during DT. To better unravel the process of DT, future research is warranted combining the analysis of the underlying processes of DT with an investigation of situational (e.g., task instructions and set-up) and individual factors (e.g., age, intelligence level, attention skills, language skills, et cetera) that may influence the involvement of different DT processes during idea generation.

Studying DT by Focusing on the Thinking Processes

Although the findings of Chapter 3 (and partly also Chapter 2) supported the idea that both memory retrieval and mental operations are effortful, top-down controlled executive processes, the executive nature of these DT processes is likely task- and process-specific. We expected involvement of or at least shared variance with domain-general EFs in DT and the DT processes. However none of the measured domain-general EFs (shifting, inhibition, and verbal working memory) appeared to be related to these DT processes, nor to predict the product measures fluency and originality over and beyond the specific DT-processes. While selective attention as a domain-general EF was found to be significantly related to fluency and originality, the effect was in the opposite direction of what was expected: low selective attention predicted higher originality, possibly indicating that a broader not-focused

attentional scope allows for discovering new, more remote associations and novel combinations of object properties (for similar findings in a study among 10-year-olds, see Van Dijk et al., 2020). Taken together, and also in view of the inconsistent findings in other studies on the relationships between DT, creativity and domain-general EFs, as reviewed in Chapter 3, these findings call for a reconceptualization of the role of domain-general EFs. While it is still plausible that a general attentional capacity (which can differ between individuals) is required for specific DT processes, the cognitive processes enabled by this general capacity are strongly dependent on the type and modality of information that is processed given the task at hand, the strategies employed by subjects (e.g., allocating the same amount of attention to either focused or broadly scanning of information) and the type of response required. This account is in line with current views on the nature of working memory as a limited attentional resource capacity embedded in long-term memory and distributed over specific information processing systems, depending on the specific task (Cowan, 2017). To the extent that standard EF tasks do not resemble DT tasks in the aforementioned respects, the shared variance can indeed be expected to be limited. Future research should investigate the executive nature of specific DT processes in more detail, employing research paradigms that directly tap into the shared limited attentional resources (e.g., by influencing the attentional load of the tasks).

Among the identified specific DT processes, the process of mental operations was found to be most strongly and consistently associated with the generation of *original* ideas, both in children as young as 4 to 6 years of age (see Chapters 2 and 3) and in adults (Forthmann et al., 2019; Gilhooly et al., 2007). As we explained in the Chapters 2 and 3, the process mental operations, among other processes, consists of: (a) the mentally singling-out or isolating of properties or parts of a structured, whole everyday object and (b) the subsequent recombination of these parts and properties into a new structured whole. These ideas align with theories on how detecting novel tool use unfolds in infants from an embodied cognition and perception-action perspective (Lockman, 2000; Lockman & Kahrs, 2017; Oudgenoeg-Paz et al., 2016). That is, infants often discover novel uses of objects following an iterative process of interactions between (a) *discovering object-affordances*, which resembles the process of singling-out object properties (e.g., the thin hairs of a toothbrush), and (b) *relating object-affordances*, which resembles the process of recombining properties into a new meaningful whole (e.g., the hair of a toothbrush matches the eyelashes of a doll given that they are similar in size, shape, and texture). This converging view on how originality or novelty emerges in different research tasks and paradigms shows that studying the process of mental operations may be a promising approach if we aim to build a comprehensive theory that can connect the relatively simple, curiosity-driven exploratory actions (cf. Hoicka et al., 2016, 2018; Oudgenoeg et al., 2016) and discovery of novel tool-use of young children (cf. Lockman, 2000; Lockman & Kahrs, 2017) to the complex inventions and innovative ideas that adults can create.

Unlike mental operations, the role of memory retrieval in the generation of original ideas is still less clear-cut. Memory retrieval was associated with originality in children of 5 to 6 years of age (Chapter 3) and adults (Gilhooly et al., 2007) but not in children aged 4 years (see Chapter 2). Combining these findings with our observations during the tests, we speculate that the increasing reliance on detailed, unique personal memories in older children might be responsible for the increasing originality of ideas generated based on memory retrieval. Future research could conduct an elaborate content analysis of children's explanations and try to distinguish different types of memories as mentioned by children and investigate their relations to the generation of original ideas, in order to better define and improve the coding of this process.

Finally, future research can also consider to combine the use of product and process measures in assessing people's level of DT or creativity in order to achieve better validity, given that subjects can come up with the same products via different thinking processes. Our findings provided preliminary evidence for the validity and reliability of using process measures such as mental operations and memory retrieval: (1) There were individual differences pertaining to both the overall frequencies and the timing of applying these processes during idea generation for a particular stimulus (Chapters 2 and 3). (2) These processes underlie fluency and originality, which are commonly used to assess subjects' level of DT or creativity in the field (Chapters 2 and 3). And (3) the application of these processes showed a developmental trend with age (Chapter 4). Yet, a more thorough investigation is needed of the validity of the DT process measures in different age groups and under different testing conditions.

Practical Implications

There is increasing consensus that creativity should be included as a core objective in education (Chinese Ministry of Education, 2012; Tickell, 2011; SLO, 2020). Early childhood education in daycare centers and kindergartens is a particularly important context in this regard, for a number of reasons. First, young children, compared to older children and adults, still show a relatively high level of creativity-related characteristics such as flexibility and curiosity (Glăveanu, 2011). It is important to foster a strong fundament of creative-divergent thinking based on children's natural creativity in early childhood. Second, the predominantly child-centered pedagogy in most early childhood education settings offers more space and possibilities for the development of creativity than formal education (Doddington & Hilton, 2007). Finally, early childhood education has often been found to have a long-term impact on so called "soft" skills and competences such as creativity to the benefit of individuals and society (Heckman & Kautz, 2012; Runco et al., 2010).

The study findings reported in this dissertation also hold implications for educational practices to foster DT and, more generally, creativity in young children. The implication of the study reported in Chapter 5 is relatively straightforward. The structured LTT-P program can be directly applied in the early education field, although some adjustments and further

research would be needed prior to that. First, the LTT-P was designed to fit in with the context of the Chinese education system, regarding particularly the usual classroom setting, predominant pedagogy, and common curriculum contents (Chinese Ministry of Education, 2012). Therefore, to be applied to kindergartens in other education systems, a prior analysis of the educational context ought to be made, and the activities should be adapted accordingly. For instance, in China kindergarten education is often organized in three grade levels and children of different ages are allocated to different levels. In LTT-P we designed separate series of activities for children of the middle and senior levels, taking into account the level of cognitive, motor, and language development of children at different ages (Chinese Ministry of Education, 2012). In this regard, a thorough adaptation of the LTT-P program would be needed when it is to be used in kindergartens where children are placed in mixed-age groups, as is the case in the Netherlands and many other countries. Second, although the field study has shown that LTT-P has the potential to promote young children's creativity, more research is needed to evaluate the effectiveness of LTT-P in the future, particularly in contexts other than the Chinese education system. This comes from our reasoning that the observed effectiveness of the LTT-P might be partially specific to Chinese educational context, pertaining particularly to originality. As we have discussed in Chapter 5, children's original thinking is more likely to be hindered or under-stimulated within Chinese kindergarten education (Niu et al., 2007; Niu & Sternberg, 2001, 2003; Wu, 2004; Wu & Albanese, 2010; Zhu & Zhang, 2008). In this regard, once they were explicitly encouraged to generate, express, and share novel ideas, as promoted in LTT-P, children may show rapid progress in the development of being able to generate original responses. However, what the effects of LTT-P are in other educational contexts is not clear and requires future research.

The studies reported in Chapters 2, 3, and 4 touched upon more fundamental topics regarding the process of DT and how DT develops in early childhood. Albeit indirectly, findings reported in these studies also hold educational implications, as was already briefly addressed in the General Introduction (Chapter 1). According to Vygotsky (1978), there is a distance between children's actual developmental level and their level of potential development, which is named "the zone of proximal development". Following Vygotsky's social-constructivist theory, the best way to facilitate children's development is to provide guidance and support that match the individual zone of proximal development of each child. In our studies, we have distinguished and conceptualized different types of thinking processes involved in DT and provided detailed descriptions and reasoning regarding how these processes might facilitate the generation of original ideas. This information could be used as a basic framework for teachers to better understand children's DT and creative behavior in daily interactions, also pertaining to the inner thinking processes which may differ from child to child to a great extent. This could help teachers to accurately assess children's actual developmental level in DT and creative thinking, serving as a first step for the development of more adaptive, personalized, and differentiated teaching and instruction methods in creativity education. In addition, more concretely, a dialogical approach might be of great

value in creativity teaching. As discussed in Chapter 4, the procedure of prompting, embedded in “natural” dialogues for testing children with the AUT, seemed to have made children aged 5 to 6 to reach a DT performance level comparable to older children (Van Dijk et al., 2020) and adults (Gilhooly et al., 2007).

Another implication of the findings reported in this dissertation relates to the future design of DT and creativity promoting education programs. As shown in Chapters 2 and 3, individual differences in the frequencies of applying certain DT processes and in the timing pattern of applying particular DT processes are associated with individual differences in the generation of original ideas. Thus, one possible approach to creativity education would be designing activities that (a) familiarize children with different thinking processes and (b) help children to optimize the dynamic application of different processes during DT. Such an approach was already adopted in designing the LTT-P as reported in Chapter 5. In LTT-P, to foster children’s creativity, most activities were designed to specifically train single thinking methods that are essential to creative thinking such as observation and DT; while some activities offered opportunities for children to learn to use more than one thinking methods in completing a complex task. In a similar vein, we recommend to organize activities around the thinking processes that underlie DT in order to stimulate children’s DT. First, we observed during our studies that not all children were equally proficient in applying all identified DT processes (see Chapter 2). For some children, DT processes other than the process memory retrieval which depends on general, everyday experiences (e.g., “use a [toothbrush] to brush your teeth because I do it every day”) did not even occur once. In this regard, activities that target the training of a single DT process could be helpful in making children aware of and able to apply certain DT processes, which could serve as building blocks for the improvement of their level of DT in general. Second, as revealed by the findings reported in Chapters 2 and 3, the emergence of original ideas depends on a complex interplay of different thinking processes. In this regard, activities that stimulate children to combine the use of several DT processes in coming up with original ideas would be another critical step to take after children become skilled at using separate DT processes.

Limitations and Future Research

The limitations of each study reported in this dissertation have been described in detail in separate chapters. Here I would like to address a few of the most prominent limitations with regard to the testing of DT in children and the approach we adopted in analyzing children’s thinking reports, calling for future research.

The study design of asking children to report on their thinking processes while performing the AUT allowed us to investigate the process of DT. The first limitation relates to the verbal prompts that were used during testing to encourage children to generate more ideas and to report on their thinking processes. The use of prompts was deemed to be necessary and recommended in previous studies (Van Someren et al., 1994) given that children of this young age often experience difficulties in reporting on their mental processes

Chapter 6

if no support is given (Flavell et al., 2000; Mercer, 2008). Although the experimenters were instructed to give prompts to all children and always in a neutral and open way, they had to decide when and how many prompts were given, thus introducing subjectivity and experimenter-variance. Future research is recommended to further standardize the procedure of giving prompts, for instance, following approaches commonly adopted in learning potential tests and dynamic assessments of intellectual abilities (Resing et al., 2012; Scheele et al., 2012).

Second, even with the help of prompts, it was challenging for children to explain their thinking processes, perhaps due to the yet under-developed language competence of children this young. As a result, only for about half of the mentioned uses per study wave, a clear and codable explanation was given, leaving unclear which DT processes were underlying the uses that were not explained by the children. In addition, children's sometimes idiosyncratic utterances also presented a great challenge during coding. For instance, for some of the DT process coding categories identified in Chapter 2, it appeared difficult to reach sufficient agreement between coders. Further refinement of the child-centered dialogical approach to DT testing might be essential to support children to express their thinking more extensively and understandably. Alternatively, in addition to asking children to report on their thinking processes, future research may also carry out in-depth content analyses of mentioned uses in order to identify the semantic and thematic-episodic relationships between generated uses, the use of mental operations and environmental cues in the generation of uses, and the switch between conceptual categories of uses. Results from such studies may also provide insights into the types of thinking processes children engage in when performing DT tasks. In particular, content analyses based on children's continuous utterances in generating uses may provide insight into the process of continuous emergence of original ideas, answering questions such as whether and how the generation of an idea is facilitated or constrained by previously generated ideas, as raised in Kupers et al. (2019).

Finally, a general limitation to be noted is that our studies focused on analyzing thinking processes underlying the AUT, a standard DT task. Although empirical studies have shown that DT product measures are related to subjects' real-life creative achievement (Cramond et al., 2005; Runco et al., 2010), to what extent the DT processes we have coded from children's thinking reports on the AUT apply or relate to real-life situations where creativity emerges (e.g., creative problem solving, industrial innovations) is unclear, thus requiring further investigation.

Conclusions

The studies reported in this dissertation investigated the development of DT in children of 4 to 6 years of age, addressing three main questions: How do children come up with original ideas? How does DT develop in children? And how can we foster DT and, more generally, creativity in children? Based on the findings reported in this dissertation, we can draw a few main conclusions. From a macro-development point of view, young children are

already able to think divergently, and this ability develops rapidly in early childhood. To support creativity education in the field, a promising approach is to provide kindergartens with structured intervention programs such as the LTT-P. From a micro-development point of view, DT involves different types of thinking processes. The current dissertation shows that the processes memory retrieval and mental operations are two key processes underlying DT in young children. Memory retrieval was the most frequently mentioned process when children explained how they came up with their ideas, likely underlying mainly the fluency aspect of DT. In contrast, the process mental operations, although infrequently mentioned, was consistently associated with the generation of original ideas within DT. Finally, in addition to the processes memory retrieval and mental operations, the generation of original ideas seems to depend on a broader attentional scope, particularly in a later stage of idea generation. This dissertation illustrates how we can better understand the nature and development of DT and, more generally, creative thinking through analyzing the underlying thinking processes.

References

- Abdullah Mirzaie, R., Hamidi, F., & Anaraki, A. A. (2009). Study on the effect of science activities on fostering creativity in preschool children. *Journal of Turkish Science Education*, 3, 81–90. Retrieved from <https://www.pegem.net/dosyalar/dokuman/124752-20110902104524-8.pdf>
- Alfonso-Benlliure, V., Meléndez, J. C., & García-Ballesteros, M. (2013). Evaluation of a creativity intervention program for preschoolers. *Thinking Skills and Creativity*, 10, 112–120. <http://dx.doi.org/10.1016/j.tsc.2013.07.005>
- Ananiadou, K. & M. Claro. (2009). 21st century skills and competences for new millennium learners in OECD countries. *OECD Education Working Papers*, 41, 4–33. <http://dx.doi.org/10.1787/218525261154>
- Bai, H., Mulder, H., Moerbeek, M., Kroesbergen, E. H., & Leseman, P. P. M. (2021). Divergent thinking in four-year-old children: An analysis of thinking processes in performing the Alternative Uses Task. *Thinking Skills and Creativity*, 40, 100814. <https://doi.org/10.1016/j.tsc.2021.100814>
- Bai, H., Leseman, P. P. M., Moerbeek, M., Kroesbergen, E. H., & Mulder, H. (2021). Serial order effect in divergent thinking in five- to six-year-olds: Individual differences as related to executive functions. *Journal of Intelligence*, 9(2), 20. <https://doi.org/10.3390/jintelligence9020020>
- 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. <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., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences*, 20(2), 87–95. <https://doi.org/10.1016/j.tics.2015.10.004>
- 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*, 6(4), 309–319. <https://doi.org/10.1037/a0029171>
- 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 and Cognition*, 42(7), 1186–1197. <https://doi.org/10.3758/s13421-014-0428-8>
- Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive function in preschool children: Test-retest reliability. *Journal of Cognition and Development*, 12(2), 169–193. <https://doi.org/10.1080/15248372.2011.563485>
- Benedek, M., & Fink, A. (2019). Toward a neurocognitive framework of creative cognition: The role of memory, attention, and cognitive control. *Current Opinion in Behavioral Sciences*, 27, 116–122. <https://doi.org/10.1016/j.cobeha.2018.11.002>
- Benedek, M., Franz, F., Heene, M., & Neubauer, A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences*, 53(4), 480–485. <https://doi.org/10.1016/j.paid.2012.04.014>
- 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>
- Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 273–281. <https://doi.org/10.1037/a0027059>
- Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick's model on creativity-related differences in associative hierarchies. Evidence for a common path to uncommon thought. *Journal of creative behavior*, 47(4), 273–289.

References

- Bijvoet-van den Berg, S., & Hoicka, E. (2014). Individual differences and age-related changes in divergent thinking in toddlers and preschoolers. *Developmental Psychology*, 50(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>
- Brown, A. S. (1973). An empirical verification of Mednick's associative theory of creativity. *Bulletin of the Psychonomic Society*, 2(6), 429–430. <https://doi.org/10.3758/BF03334439>
- 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
- Cheung, R. H. P. (2010). Designing movement activities to develop children's creativity in early childhood education. *Early Child Development and Care*, 180(3), 377–385. <https://doi.org/10.1080/03004430801931196>
- Cheung, R. H. P. (2018). Play-based creativity-fostering practices: The effects of different pedagogical approaches on the development of children's creative thinking behaviours in a Chinese preschool classroom. *Pedagogy, Culture and Society*, 26(4), 511–527. <https://doi.org/10.1080/14681366.2018.1424725>
- Chinese Ministry of Education (2012). 3–6 sui ertong xuexi yu fazhan zhinan [Early learning and development guideline for 3- to 6-year-old children]. <https://www.gov.cn/gzdt/att/att/site1/20121016/001e3741a42811e6f63f01.doc>.
- Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology*, 53(2), 82–88. <https://doi.org/10.1037/h0045461>
- Chronopoulou, E., & Riga, V. (2012). The contribution of music and movement activities to creative thinking in pre-school children. *Creative Education*, 3(2), 196–204. <https://doi.org/10.4236/ce.2012.32031>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Coney, J., & Serna, P. (1995). Creative thinking from an information perspective: New approach to Mednick's theory of associative hierarchies. *Journal of Creative Behavior*, 29(2), 109–132. <https://doi.org/10.1002/j.2162-6057.1995.tb00740.x>
- Cowan, N., & Alloway, T. (2009). Development of working memory in childhood. In M. L. Courage & N. Cowan (Eds.), *Studies in developmental psychology. The development of memory in infancy and childhood* (pp. 303–342). Psychology Press.
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychological Bulletin Review*, 24(4), 1158–1170. <https://doi.org/10.3758/s13423-016-1191-6>
- Conway, A. R. A., & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, 4(6), 577–590. <https://doi.org/10.1080/741940997>
- Craft, A., McConnon, L., & Matthews, A. (2012). Child-initiated play and professional creativity: Enabling four-year-olds' possibility thinking. *Thinking Skills and Creativity*, 7(1), 48–61. <https://doi.org/10.1016/j.tsc.2011.11.005>
- Cramond, B., Matthews-Morgan, J., Bandalos, D., & Zuo, L. (2005). A report on the 40-year follow-up of the Torrance Tests of Creative Thinking: Alive and well in the new millennium. *Gifted Child Quarterly*, 49(4), 283–291. <https://doi.org/10.1177/001698620504900402>
- Daugherty, M. (1993). Creativity and private speech: Developmental trends. *Creativity Research Journal*, 6(3), 287–296. <https://doi.org/10.1080/10400419309534484>

- De Dreu, C. K. W., Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working memory benefits creative insight, musical improvisation, and original ideation through maintained task-focused attention. *Personality and Social Psychology Bulletin*, 38(5), 656–669. <https://doi.org/10.1177/0146167211435795>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Doddington, C., & Hilton, M. (2007). Child-centred education: Reviving the creative tradition. Sage.
- Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children. *Developmental Review*, 38, 241–268. <https://doi.org/10.1016/j.dr.2015.09.001>
- Dziedziewicz, D., Oledzka, D., & Karwowski, M. (2013). Developing 4- to 6-year-old children's figural creativity using a doodle-book program. *Thinking Skills and Creativity*, 9, 85–95. <https://doi.org/10.1016/j.tsc.2012.09.004>
- 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>
- Edl, S., Benedek, M., Papousek, I., Weiss, E. M., & Fink, A. (2014). Creativity and the Stroop interference effect. *Personality and Individual Differences*, 69, 38–42. <https://doi.org/10.1016/j.paid.2014.05.009>
- Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology*, 62, 559–582. <https://doi.org/10.1146/annurev.psych.121208.131702>
- 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(1), 97–112. https://doi.org/10.1207/S15327647JCD0101N_10
- Forthmann, B., Szardenings, C., & Holling, H. (2018). Understanding the confounding effect of fluency in divergent thinking scores: Revisiting average scores to quantify artifactual correlation. *Psychology of Aesthetics, Creativity, and the Arts*, 14(1), 94–112. <https://doi.org/10.1037/aca0000196>
- Forthmann, B., Wilken, A., Doebler, P., & Holling, H. (2019). Strategy induction enhances creativity in figural divergent thinking. *Journal of Creative Behavior*, 53(1), 18–29. <https://doi.org/10.1002/jocb.159>
- Gao, H., & Huo, L. (2017). The Early Learning and Development Guidelines in China: Retrospective and prospective views. In M. Li, J. Fox, & S. Grieshaber. (Eds), *Contemporary Issues and Challenge in Early Childhood Education in the Asia-Pacific Region* (pp. 181–197). Springer.
- Garaigordobil, M., & Berruero, L. (2011). Effects of a play program on creative thinking of preschool children. *The Spanish Journal of Psychology*, 14(2), 608–618.
- Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 39(1), 3–27. <https://doi.org/10.1017/S0021963097001753>
- 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>
- Glăveanu, V. P. (2011). Children and creativity: A most (un)likely pair? *Thinking Skills and Creativity*, 6(2), 122–131. <https://doi.org/10.1016/j.tsc.2011.03.002>
- Glăveanu, V. P. (2013). Rewriting the language of creativity: The five A's framework. *Review of General Psychology*, 17(1), 69–81. <https://doi.org/10.1037/a0029528>
- Glăveanu, V. P. (2018). Creativity in perspective: A sociocultural and critical account. *Journal of Constructivist Psychology*, 31(2), 118–129. <https://doi.org/10.1080/10720537.2016.1271376>

References

- Glăveanu, V. P., Hanchett Hanson, M., Baer, J., Barbot, B., Clapp, E. P., Corazza, G. E., Hennessey, B., Kaufman, J. C., Lebuda, I., Lubart, T., Montuori, A., Ness, I. J., Plucker, J., Reiter-Palmon, R., Sierra, Z., Simonton, D. K., Neves-Pereira, M. S. & Sternberg, R.J. (2020). Advancing creativity theory and research: A socio-cultural manifesto. *Journal of Creative Behavior*, 54(3), 741–745. <https://doi.org/10.1002/jocb.395>
- 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*, 3(1), 152–177. <https://doi.org/10.1515/ctra-2016-0011>
- Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, 53(4), 267–293. <https://doi.org/10.1037/h0040755>
- Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill Book Co.
- Gupta, A. (2008). Vygotskian perspectives on using dramatic play to enhance children’s development and balance creativity with structure in the early childhood classroom. *Early Child Development and Care*, 8, 1041–1054. <https://doi.org/10.1080/03004430701731654>.
- 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). Semantic search during divergent thinking. *Cognition*, 166, 344–357. <https://doi.org/10.1016/j.cognition.2017.05.039>
- Hass, R. W. (2017b). Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. *Memory and Cognition*, 45(2), 233–244. <https://doi.org/10.3758/s13421-016-0659-y>
- Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451–464. <https://doi.org/10.1016/j.labeco.2012.05.014>
- Hedeker, D., Gibbons, R., du Toit, M., & Cheng, Y. (2008). *SuperMix: Mixed effects models*. Scientific Software International. https://ssicentral.com/wp-content/uploads/2020/07/SMIX_Manual.pdf
- 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>
- Hendry, A., Jones, E. J. H., & Charman, T. (2016). Executive function in the first three years of life: Precursors, predictors and patterns. *Developmental Review*, 42, 1–33. <https://doi.org/10.1016/j.dr.2016.06.005>
- Hoff, E. (2009). *Language development* (4th ed.). Michele Sordi.
- Hoicka, E., Bijvoet-Van Den Berg, S., Kerr, T., & Carberry, M. (2013). The Unusual Box Test: A non-verbal, non-representational divergent thinking test for toddlers. www.aaai.org
- 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*, 87(4), 1099–1105. <https://doi.org/10.1111/cdev.12531>
- Hoicka, E., Powell, S., Knight, J., & Norwood, M. (2018). Two-year-olds can socially learn to think divergently. *British Journal of Developmental Psychology*, 36(1), 22–36. <https://doi.org/10.1111/bjdp.12199>
- Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.
- Hu, W., Adey, P., Jia, X., Liu, J., Zhang, L., Li, J., & Dong, X. (2011). Effects of a “Learn to Think” intervention programme on primary school students. *British Journal of Educational Psychology*, 81(4), 531–557. <https://doi.org/10.1348/2044-8279.002007>

- Hu, W., Jia, X., Liu, J., Shan, X., & Boyanton, D. (2016). Effects of a “Learn to Think” intervention program on Chinese primary school students’ learning strategies. *The International Journal of Creativity and Problem Solving*, 26(1), 21–42.
- Hu, W., Jia, X., Plucker, J.A., & Shan, X. (2016). Effects of a critical thinking skills program on the learning motivation of primary school students. *Roeper Review*, 38, 1–14.
<https://doi.org/10.1080/02783193.2016.1150374>.
- Hu, W., Wu, B., Jia, X., Yi, X., Duan, C., Meyer, W., & Kaufman, J. C. (2013). Increasing students’ scientific creativity: The “Learn to Think” intervention program. *Journal of Creative Behavior*, 47(1), 3–21.
<https://doi.org/10.1002/jocb.20>
- Jamali, B., Mohammad Kazemi, R., & Shahbazi, M. (2012). Effects of sport activities on increasing preschool children’s creativity. *Management Science Letters*, 2(6), 1975–1980.
<https://doi.org/10.5267/j.msl.2012.06.024>
- Johns, G. (2017). Reflections on the 2016 decade award: Incorporating context in organizational research. *Academy of Management Review*, 4, 577–595. <http://dx.doi.org/10.5465/amr.2017.0044Invited>
- Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development*, 45(1), 65–94.
<https://doi.org/10.1007/bf02299613>
- Karr, J. E., Areshenkov, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, 144(11), 1147–1185. <https://doi.org/10.1037/bul0000160>
- Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, 13(1), 1–12. <https://doi.org/10.1037/a0013688>
- Kharkhurin, A. v., & Wei, L. (2015). The role of code-switching in bilingual creativity. *International Journal of Bilingual Education and Bilingualism*, 18(2), 153–169. <https://doi.org/10.1080/13670050.2014.884211>
- Kim, K. H. (2011). The creativity crisis: The decrease in creative thinking scores on the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 23(4), 285–295. <https://doi.org/10.1080/10400419.2011.627805>
- 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*, 8(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>
- Kupers, E., Lehmann-Wermser, A., McPherson, G., & Van Geert, P. (2019). Children’s creativity: A theoretical framework and systematic review. *Review of Educational Research*, 89(1), 93–124.
<https://doi.org/10.3102/0034654318815707>
- Lee, C. S., & Theriault, 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*, 41(5), 306–320. <https://doi.org/10.1016/j.intell.2013.04.008>
- Lin, Y. S. (2010). Drama and possibility thinking - Taiwanese pupils’ perspectives regarding creative pedagogy in drama. *Thinking Skills and Creativity*, 5(3), 108–119. <https://doi.org/10.1016/j.tsc.2010.09.001>
- Lockman, J. J. (2000). A perception-action perspective on tool use development. *Child Development*, 71(1), 137–144. <https://doi.org/10.1111/1467-8624.00127>

References

- Lockman, J. J., & Kahrs, B. A. (2017). New insights into the development of human tool use. *Current Directions in Psychological Science*, 26(4), 330–334. <https://doi.org/10.1177/0963721417692035>
- Lu, J. G., Akinola, M., & Mason, M. F. (2017). “Switching on” creativity: Task switching can increase creativity by reducing cognitive fixation. *Organizational Behavior and Human Decision Processes*, 139, 63–75. <https://doi.org/10.1016/j.obhdp.2017.01.005>
- Luo, M., & Li, K. (2017). Jiyu quanguo 428 ge banji yangben de xueqian jiaoyu zhiliang chengxiang chaju toushi [Reviewing the urban-rural gap regarding the quality of early education: Based on a national sample of 428 classes]. *Studies in Early Childhood Education*, 6, 13–20.
- Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220–232. <https://doi.org/10.1037/h0048850>
- Mercer, N. (2008). Talk and the development of reasoning and understanding. *Human Development*, 51(1), 90–100. <https://doi.org/10.1159/000113158>
- Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.
- Mertler, C. A., & Vannatta, R. A. (2013). *Advanced and multivariate statistical methods: Practical Application and Interpretation*. Routledge.
- Milgram, R. M., & Rabkin, L. (1980). Developmental test of Mednick's associative hierarchies of original thinking. *Developmental Psychology*, 16(2), 157–158. <https://doi.org/10.1037/0012-1649.16.2.157>
- Mirzaie, R. A., Hamidi, F., & Anaraki, A. (2009). A study on the effect of science activities on fostering creativity in preschool children. *Journal of Turkish Science Education*, 6(3), 81–90.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8–14. <https://doi.org/10.1177/0963721411429458>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Moriguchi, Y., Chevalier, N., & Zelazo, P. D. (2016). Editorial: Development of executive function during childhood. *Frontiers in Psychology*, 7(6), 1–2. <https://doi.org/10.3389/fpsyg.2016.00006>
- Mulder, H., Hoofs, H., Verhagen, J., Van der Veen, I., & Leseman, P. P. M. (2014). Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds. *Frontiers in Psychology*, 5, 1–17. <https://doi.org/10.3389/fpsyg.2014.00733>
- Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, 111(2). <https://doi.org/10.1037/0033-295X.111.2.486>
- Newcombe, N. S. (2002). *Spatial cognition*. In H. Pashler & D. Medin (Eds.), *Steven's handbook of experimental psychology: Memory and cognitive processes* (pp. 113–163). John Wiley & Sons Inc.
- Newman, J. P., Widom, C. S., & Nathan, S. (1985). Passive avoidance in syndromes of disinhibition. Psychopathy and Extraversion. *Journal of Personality and Social Psychology*, 48(5), 1316–1327. <https://doi.org/10.1037/0022-3514.48.5.1316>
- Nguyen, R., Brooks, M., Bruno, R., & Peacock, A. (2018). Behavioral measures of state impulsivity and their psychometric properties: A systematic review. *Personality and Individual Differences*, 135, 67–79. <https://doi.org/10.1016/j.paid.2018.06.040>
- Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, 21(1), 34–77. <https://doi.org/10.1080/10463281003765323>
- Niu, W., & Sternberg, R. J. (2001). Cultural influences on artistic creativity and its evaluation. *International Journal of Psychology*, 36(4), 225–241. <https://doi.org/10.1080/002075901430000036>

- Niu, W., & Sternberg, R. J. (2003). Societal and school influences on student creativity: The case of China. *Psychology in the Schools*, 40(1), 103–114. <https://doi.org/10.1002/pits.10072>
- Niu, W., Zhang, J. X., & Yang, Y. (2007). Deductive reasoning and creativity: A cross-cultural study. *Psychological Reports*, 100(2), 509–519. <https://doi.org/10.2466/PRO.100.2.509-519>
- 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*, 39(1), 36–45. <https://doi.org/10.1016/j.intell.2010.11.002>
- Oudgenoeg-Paz, O., Volman, M. C., & Leseman, P. P. M. (2016). First steps into language? Examining the specific longitudinal relations between walking, exploration and linguistic skills. *Frontiers in psychology*, 7, 1458. <https://doi.org/10.3389/fpsyg.2016.01458>
- Piirto, J. (2011). *Creativity for 21st century skills: How to embed creativity into the curriculum*. Sense Publishers. https://doi.org/10.1007/978-94-6091-463-8_1
- Plucker, J.A., Beghetto, R.A., & Dow, G.T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, 2, 83–96. https://doi.org/10.1207/s15326985ep3902_1
- Pringle, A., & Sowden, P. T. (2017). The Mode Shifting Index (MSI): A new measure of the creative thinking skill of shifting between associative and analytic thinking. *Thinking Skills and Creativity*, 23, 17–28. <https://doi.org/10.1016/j.tsc.2016.10.010>
- 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>
- Reiter-Palmon, R., Forthmann, B., & Barbot, B. (2019). Scoring divergent thinking tests: A review and systematic framework. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 144–152. <https://doi.org/10.1037/aca0000227>
- Resing, W. C. M., Elliott, J. G., & Grigorenko, E. L. (2012). Dynamic testing and assessment. In N. M. Seel (Eds), *Encyclopedia of the Sciences of Learning* (Vol. 50, pp. 1055–1058). Springer. <https://doi.org/10.1007/978-1-4419-1428-6>
- Riga, V., & Chronopoulou, E. (2014). Applying MacKinnon's 4Ps to foster creative thinking and creative behaviours in kindergarten children. *Education 3–13*, 42(3), 330–345. <https://doi.org/10.1080/03004279.2012.692700>
- Rizi, C. E., Yarmohamadian, M. H., & Gholami, A. (2011). The effect group plays on the development of the creativity of six-year children. *Procedia - Social and Behavioral Sciences*, 15, 2137–2141. <https://doi.org/10.1016/j.sbspro.2011.04.067>
- Robson, S. (2014). The Analysing Children's Creative Thinking framework: Development of an observation-led approach to identifying and analysing young children's creative thinking. *British Educational Research Journal*, 40(1), 121–134. <https://doi.org/10.1002/berj.3033>
- Rominger, C., Fink, A., Weiss, E. M., Bosch, J., & Papousek, I. (2017). Allusive thinking (remote associations) and auditory top-down inhibition skills differentially predict creativity and positive schizotypy. *Cognitive Neuropsychiatry*, 22(2), 108–121. <https://doi.org/10.1080/13546805.2016.1278361>
- Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, 24(1), 66–75. <https://doi.org/10.1080/10400419.2012.652929>
- Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92–96. <https://doi.org/10.1080/10400419.2012.650092>
- Runco, M. A., Millar, G., Acar, S., & Cramond, B. (2010). Torrance Tests of Creative Thinking as predictors of personal and public achievement: A fifty-year follow-up. *Creativity Research Journal*, 22(4), 361–368. <https://doi.org/10.1080/10400419.2010.523393>

References

- Russ, S. W., & Wallace, C. E. (2013). Pretend play and creative processes. *American Journal of Play*, 6, 136–148. <https://doi.org/10.1097/SCS.0b013e318262db2f>
- 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*, 112(3), 419–444. <https://doi.org/10.1086/663300>
- Scibinetti, P., Tocci, N., & Pesce, C. (2011). Motor creativity and creative thinking in children: The diverging role of inhibition. *Creativity Research Journal*, 23(3), 262–272. <https://doi.org/10.1080/10400419.2011.595993>
- 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. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, 44(4), 1012–1021. <https://doi.org/10.1016/j.paid.2007.10.027>
- Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychology Review*, 27(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. *Journal of Creative Behavior*, 4, 345–347. <https://doi.org/10.1002/jocb.194>
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (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(2), 68–85. <https://doi.org/10.1037/1931-3896.2.2.68>
- Simonton, D. K. (2012). Taking the U.S. patent office criteria seriously: A quantitative three-criterion creativity definition and its implications. *Creativity Research Journal*, 24(2–3), 97–106. <https://doi.org/10.1080/10400419.2012.676974>
- SLO. (2020, 8 December). *Spel en begeleiding met jonge kinderen* [Playing and guiding with young children]. <https://www.slo.nl/thema/meer/jonge-kind/spel-begeleiding/>
- SLO. (2020, 28 December). *Creatief denken en handelen* [Creative thinking and behavior]. <https://www.slo.nl/thema/meer/21e-eeuwsevaardigheden/creatief-denken/>
- Smith, G. J. W., & Carlsson, I. (1983). Creativity in early and middle school years. *International Journal of Behavioral Development*, 6(2), 167–195. <https://doi.org/10.1177/016502548300600204>
- Smith, J. K., & Smith, L. F. (2017). The 1.5 criterion model of creativity: Where less is more, more or less. *Journal of Creative Behavior*, 51(4), 281–284. <https://doi.org/10.1002/jocb.191>
- Smogorzewska, J. (2012). Storyline and associations pyramid as methods of creativity enhancement: Comparison of effectiveness in 5-year-old children. *Thinking Skills and Creativity*, 7, 28–37. <https://doi.org/10.1016/j.tsc.2011.12.003>
- Smogorzewska, J. (2014). Developing children's language creativity through telling stories - An experimental study. *Thinking Skills and Creativity*, 13, 20–31. <https://doi.org/10.1016/j.tsc.2014.02.005>
- Sowden, P. T., Pringle, A., & Gabora, L. (2015). The shifting sands of creative thinking: Connections to dual-process theory. *Thinking and Reasoning*, 21(1), 40–60. <https://doi.org/10.1080/13546783.2014.885464>
- Stolte, M., Kroesbergen, E. H., & Van Luit, J. E. H. (2019). Inhibition, friend or foe? Cognitive inhibition as a moderator between mathematical ability and mathematical creativity in primary school students. *Personality and Individual Differences*, 142, 196–201. <https://doi.org/10.1016/j.paid.2018.08.024>
- Tickell, C. (2011). *The early years: Foundations for life, health and learning. An independent report on the early years foundation stage to Her Majesty's government*. <http://outdoormatters.co.uk/wp-content/uploads/2011/04/The-Early-Years-Foundations-for-life-health-and-learning.pdf>

- Torrance, E. P. (1968). A longitudinal examination of the fourth grade slump in creativity. *Gifted Child Quarterly*, 4, 195–199. [https://doi.org/10.1016/0001-8791\(92\)90037-Z](https://doi.org/10.1016/0001-8791(92)90037-Z)
- Torrance, E. P. (1972). Predictive validity of the Torrance Tests of Creative Thinking. *Journal of Creative Behavior*, 6(4), 236–252. <https://doi.org/10.1002/j.2162-6057.1972.tb00936.x>
- Torrance, E.P., Ball, O.E., & Safter, H.T. (1992/2008). *Torrance tests of creative thinking. Directions manual and scoring guide for figural forms A and B*. Scholastic Testing Service.
- Urban, K. K. (1991). On the development of creativity in children. *Creativity Research Journal*, 4(2), 177–191. <https://doi.org/10.1080/10400419109534384>
- Van de Kamp, M. T., Admiraal, W., & Rijlaarsdam, G. (2016). Becoming original: Effects of strategy instruction. *Instructional Science*, 44(6), 543–566. <https://doi.org/10.1007/s11251-016-9384-y>
- Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *British Journal of Developmental Psychology*, 31(1), 70–87. <https://doi.org/10.1111/j.2044-835X.2012.02079.x>
- Van Dijk, M., Kroesbergen, E. H., Blom, E., & Leseman, P. P. M. (2019). Bilingualism and creativity: Towards a situated cognition approach. *Journal of Creative Behavior*, 53(2), 178–188. <https://doi.org/10.1002/jocb.238>
- Van Dijk, M., Blom, E., Kroesbergen, E. H., & Leseman, P. P. M. (2020). The influence of situational cues on children's creativity in an alternative uses task and the moderating effect of selective attention. *Journal of Intelligence*, 8(4), 1–20. <https://doi.org/10.3390/jintelligence8040037>
- Van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. Academic Press. [https://doi.org/10.1016/0306-4573\(95\)90031-4](https://doi.org/10.1016/0306-4573(95)90031-4)
- Veer, I. M., Luyten, H., Mulder, H., Van Tuijl, C., & Slegers, P. J. C. (2017). Selective attention relates to the development of executive functions in 2,5- to 3-year-olds: A longitudinal study. *Early Childhood Research Quarterly*, 41, 84–94. <https://doi.org/10.1016/j.ecresq.2017.06.005>
- Verhagen, J., Boom, J., Mulder, H., De Bree, E., & Leseman, P.P.M. (2019). Reciprocal relationships between nonword repetition and vocabulary during the preschool years. *Developmental Psychology*, 55(6), 1125–1137. <http://dx.doi.org/10.1037/dev0000702>
- Verhagen, J., & Mulder, H. (2013). *Testinstructie voor de testleiders van het cohortonderzoek Pre-COOL: Period 2013–2014* [Test instructions for the experimenters of the cross-sectional research Pre-COOL: Period 2013–2014][Unpublished manuscript]. Department of Pedagogy and Education, Development & Education of Youth in Diverse Societies, Utrecht University.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological process*. Harvard University Press.
- 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>
- Wang, L., & Chen, Z. (2014). 3–6 sui you'er qinziran qinggan de peiyang celue [Strategies in cultivating 3- to 6-year-old children's emotional affinity towards nature]. *Studies in Early Childhood Education*, 6, 61–63.
- Ward, W. C. (1969). Rate and uniqueness in children's creative responding. *Child Development*, 40(3), 869–878. <https://doi.org/10.1111/j.1467-8624.1969.tb04548.x>
- Weisberg, R. W. (2015). On the usefulness of “value” in the definition of creativity. *Creativity Research Journal*, 27(2), 111–124. <https://doi.org/10.1080/10400419.2015.1030320>
- 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>

References

- Wong, R., & Niu, W. (2013). Cultural difference in stereotype perceptions and performances in nonverbal deductive reasoning and creativity. *Journal of Creative Behavior*, 47 (1), 41–59. <http://dx.doi.org/10.1002/jocb.22>
- Wu, J.-J. (2004). Recognizing and nurturing creativity in Chinese students. In S. Lau, A. Hui & G. Ng (Eds.), *Creativity: When east meets west* (pp. 169–200). World Scientific Publishing Pvt. Ltd.
- Wu, J.-J., & Albanese, D. (2010). Asian creativity, chapter one: Creativity across three Chinese societies. *Thinking Skills and Creativity*, 5, 150–154. <http://dx.doi.org/10.1016/j.tsc.2010.10.002>
- Ye, P., & Ma, Q. (2012). 2–6 sui ertong chuangzaoxing siwei fazhang de tedian ji guilv [The characteristics and laws of the development of creative thinking of children aged 2–6 years old]. *Studies in Early Childhood Education*, 81, 36–41.
- Yeh, Y. C., Lai, G. J., Lin, C. F., Lin, C. W., & Sun, H. C. (2015). How stress influences creativity in game-based situations: Analysis of stress hormones, negative emotions, and working memory. *Computers and Education*, 81, 143–153. <https://doi.org/10.1016/j.compedu.2014.09.011>
- Zabelina, D. L. (2018). Attention and creativity. In *The Cambridge Handbook of the Neuroscience of Creativity* (pp. 161–179). Cambridge University Press. <https://doi.org/10.1017/9781316556238.010>
- Zabelina, D. L., Friedman, N. P., & Andrews-Hanna, J. (2019). Unity and diversity of executive functions in creativity. *Consciousness and Cognition*, 68, 47–56. <https://doi.org/10.1016/j.concog.2018.12.005>
- Zabelina, D. L., & Ganis, G. (2018). Creativity and cognitive control: Behavioral and ERP evidence that divergent thinking, but not real-life creative achievement, relates to better cognitive control. *Neuropsychologia*, 118(A), 20–28. <https://doi.org/10.1016/j.neuropsychologia.2018.02.014>
- Zabelina, D. L., & Robinson, M. D. (2010). Creativity as flexible cognitive control. *Psychology of Aesthetics, Creativity, and the Arts*, 4(3), 136–143. <https://doi.org/10.1037/a0017379>
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297–301. <https://doi.org/10.1038/nprot.2006.46>
- Zhu, J., & Zhang, J. (2008). Contemporary trends and developments in early childhood education in China. *Early Years*, 2, 173–182. <https://doi.org/10.1080/09575140802163584>

Appendices

Appendix 2.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 2.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.

Appendices

Appendix 3.A

Main Types of Actions Used for Classifying Uses and Scoring Their Originality

No.	Lunch box	Tire	Pencil	Umbrella	Shovel	Toothbrush
1	To put food in / as a food box	For cars / to ride with	For drawing	For (/ to protect you from) rain(ing)	To shovel/dig	To clean teeth
2	To put something in / as a storage box	For making a swing	For writing	As a boat	To beat something or someone / to break something	To clean something / as a cleaning brush
3	To put insect or animal in / as an insect box	To play with (unspecified)	To erase something (use the attached eraser)	As a sunshade / against the sun	To put something on or in	To clean someone
4	For cut-and-glue activity / for artwork / to make something (unspecified)	For sitting or lying on	For coloring	To put something in	For fighting / as a sword	For painting or drawing
5	To put water or drink in	To sit in and roll	To poke something / to break something	For flying (with it)	To build or make something with sand	To sweep something / as a broom
6	As a garbage bin / to put dirty things in	To roll with (a sort of game)	For scratching	For beating something	To pick up or lift something	To brush hair / as a comb
7	For catching something	As a hula hoop	For cut-and-glue activity / to make something (unspecified)	For fighting / as a sword	For planting	For cut-and-glue activity / to make something (unspecified)
8	For taking something with you	For jumping in	For beating something	For protecting yourself / as a shield	For building or construction work	To play with (unspecified)
9	To write/draw something on	For making a small house	To draw something for cut-and-glue activity	To write/draw something on	For looking for something (usually in sand)	For fighting / as a sword
10	For making a boat	To put something in and roll	For painting	As a merry-go-round/carousel	To chop or hack something	For pretending to brush your teeth.

Note. Here we only present action types that were mentioned by at least five children (5% of the 101 children)

Appendix 3.B

Inter-Coder Reliability Overview and Kappa's for Divergent Thinking (DT) Processes (399 episodes)

DT process	Frequency of coded episodes by the two coders				Kappa's
	yes-yes	yes-no	no-yes	no-no	
Memory retrieval	151	25	14	209	.80
Mental operations	34	21	5	339	.69

Appendix 3.C

Missing Data of Different Tasks and Reasons for the Missing

Task		Final sample size	Missing data and the reasons
Divergent thinking	AUT	101	1 child was not present on test dates;
Inhibition	Go/NoGo	100	1 child was simply hitting the key all the time; 1 child's data file was overwritten by mistake.
	Animal stroop	100	1 child was sick on test dates; 1 child's test session was interrupted.
Shifting	Animal shifting	102	No missing data for this task.
	DCCS	96	1 child's recording form was lost; 5 children were not tested following the instructions from the manual (e.g. the reminders of rules were not given while a stimulus was presented to children).
Working memory	Word recall	98	4 children showed clear indications of not understanding the instructions (i.e., try to recall the words in reversed alphabet order)
Selective attention	Elephant search	101	1 child was sick on test dates.

Note. AUT = The Alternative Uses Task; DCCS = The Dimensional Change Card Sort task

Appendices

Appendix 4.A

Multilevel Regression Analyses: Examining the Testing Effect (N = 106, with 287 units at the measurement level)

Model	Intercept-only model		The effect of group		Intercept-only model		The effect of group	
	Fluency	Originality	Fluency	Originality	Memory retrieval	Mental operations	Memory retrieval	Mental operations
Fixed part	Coeff. (s.e.)							
Intercept	3.73 (0.12)***	1.69 (0.08)***	3.80 (0.16)***	1.64 (0.11)***	1.41 (0.06)***	0.34 (0.04)***	1.39 (0.09)***	0.29 (0.06)***
Age			1.04 (0.18)***	0.89 (0.12)***			0.54 (0.10)***	0.21 (0.07)**
Group			-0.30 (0.23)	-0.02 (0.15)			-0.02 (0.12)	0.06 (0.08)
Age × group			-0.07 (0.26)	-0.03 (0.18)			-0.09 (0.14)	0.14 (0.10)
Random slope								
σ^2_{age}			0.12 (0.29)	0.18 (0.13)			0.11 (0.08)	0.05 (0.04)
Random effect								
σ^2_{child}	0.56 (0.22)*	0.20 (0.09)*	0.71 (0.21)***	0.35 (0.09)**	0.15 (0.06)**	0.08 (0.03)**	0.24 (0.06)***	0.11 (0.03)**
σ^2_{wave}	2.44 (0.26)***	1.13 (0.12)***	1.81 (0.24)***	0.66 (0.09)***	0.64 (0.07)***	0.24 (0.03)***	0.42 (0.06)***	0.20 (0.03)***
Deviance (df)	1632.50 (8)		1493.55 (21)		1189.88 (8)		1065.31 (21)	
Δ deviance (Δ df), vs. the previous model			138.95 (13)***				124.57 (13)***	

Note. All models were run with data of the first, third, and fourth measurement waves in order to examine whether the testing effect was present.

⁺ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Appendix 5.A

Training Goals of Each Thinking Method for Different Grade Levels

Thinking Method	Middle level	High level
Observation	<p>Observe the typical features of objects under teachers' instructions.</p> <p>Observe and perceive objects or events in order (e.g., according the alignment, time, or other rules) under teachers' instructions.</p> <p>Observe and compare different visualized materials to find the differences.</p>	<p>Find the features of objects through observation.</p> <p>Observe objects in order and start to understand the strategy of observing in order under teachers' instructions.</p> <p>Conduct systematic observation based on visualized materials and find the similarities and differences.</p>
Association	<p>Associate different operative or visualized materials based on the similarities.</p> <p>Associate described scenarios to related objects or themes.</p>	<p>Skilled at associating visualized materials or objects/things based on similarities or other relations within familiar scenarios.</p> <p>Compare and associate the features/properties of different operative/visualized materials under teachers' instructions.</p>
Imagination	<p>Understand task requirements.</p> <p>Build up imagination based on established scenarios or visualized materials under teachers' instructions.</p> <p>Imagine with intention.</p>	<p>Skilled at building up imagination based on established scenarios or visualized materials.</p> <p>Start to imagine something with novelty and try to think about the possibility of realizing the imagination under teachers' instructions.</p>
Aligning	<p>Find differences among and compare the same type of operative or visual objects.</p> <p>Use established standards to align objects.</p>	
Classification	<p>Understand the basic idea of "classification".</p> <p>Distinguish different uses of common objects based on daily experiences.</p>	<p>Understand the concept of "classification".</p> <p>Classify different objects based on certain features such as color, size, and function.</p>
Reasoning	<p>Do simple reasoning with established stories or other established scenarios (with one cue) under teachers' instructions.</p>	<p>Do reasoning with established stories or other complex established scenarios (with two or three cues) under teachers' instructions.</p>
Reconstruction	<p>Understand the basic idea of "reconstruction".</p> <p>Reconstruct new things with provided materials and describe what has been made.</p>	<p>Discover how certain objects or things are reconstructed by other more basic, elementary things.</p> <p>Use provided materials to reconstruct new things which contains some extent of novelty under teachers' instructions; and explain to what situations can the new things be applied and how the idea comes up.</p>

Appendices

Thinking Method	Middle level	High level
Divergent thinking	<p>Explore an object or a scenario from different aspects.</p> <p>Develop simple strategies that help promoting divergent thinking under teachers' instructions (e.g., think about different environments)</p>	<p>Think divergently with an established question.</p> <p>Apply simple divergent thinking strategies (e.g., think about different environment, decompose the objects) to explore a question and generate possible solutions under teachers' instructions.</p> <p>Express different ideas through drawing, movement, or language.</p>
Artistic creation	<p>Express ideas with artistic outcomes under teachers' instructions.</p> <p>Children are given spaces and allowed to do replicative creation.</p> <p>It is more important about the process of expressing and creating, rather than the quality of the final products.</p>	<p>Create artistic products with richer content that have some extent of novelty.</p> <p>Not only the process of expressing ideas, but also the reflection about the process and the products are emphasized.</p> <p>Novelty are expected in both aspects.</p>
Problem finding	<p>Ask simple questions (e.g., "What is it?", "How does it work", etc.) about daily phenomena or (artificially) established scenarios under teachers' instructions.</p>	<p>Ask complex, exploratory questions (e.g., 'Why is it work like this?', 'Why is it different that...?', etc.) about natural phenomena or (artificially) established scenarios under teachers' instructions.</p> <p>Children are expected to combine several basic thinking methods, such as observation, comparing, mindset breaking, to form informative questions.</p>
Scientific Exploration	<p>Become interested in scientific exploration</p>	<p>Find questions and describe them.</p> <p>Think about possible solutions for certain questions under teachers' instructions</p>

Appendix 5.B

Different Activities Distributed Across Thinking Methods and Content Domains

Thinking Method	Music & Movement	Language	Graphic /Drawing	Nature	Math & Geometry
Observation	Hair shoulder knees toe Hide-and-seek of the shadow			My little precious*	Animal puzzles The world of puzzle*
Association	The picture of music*		The network of an apple The family of circles		My own supermarket*
Imagination	If I were an animal		Finger drawing The upside- down world*		
Aligning					Line up the bottles
Classification				The traveling of the seeds*	Find a home for toys
Reasoning	Challenge 1 2 3	Junior detectors Save the squirrel*			Graphic reasoning*
Reconstruction			The secrets of figures I am an inventor!*		The tangram with thousands of forms
Divergent thinking		How to play with a grapefruit Magic world of chopsticks Special plastic bags* Unexpected uses of newspapers*	New uses of the hat The show of squares*	How do the chicken go through the river?*	
Artistic creation		The little black fish What happened in the fruit garden?*	The most lovely scarf Ceremony for fruits* Collage with beans*		
Problem finding		The king of storytelling* Able to ask questions?		Ten thousand questions about the marine*	
Scientific exploration				The paper boat never drowns The compass*	

*These activities are for children at the senior level, and the rest are for children at the middle level.

Samenvatting

(Summary in Dutch)

Divergent denken (DD), gedefinieerd als het denkproces dat mensen in staat stelt om een verscheidenheid aan ideeën te genereren, krijgt steeds meer aandacht en er is redelijk veel onderzoek naar gedaan in de afgelopen decennia. DD is, als belangrijke component van creativiteit, van groot belang voor het bedenken van originele of nieuwe ideeën en wordt gezien als een belangrijke factor in innovatie en probleemoplossing. In de wetenschappelijke literatuur worden verschillende tests beschreven om het niveau van DD in een individu te meten. Dergelijke tests zijn gebruikt om onderzoek te doen naar de ontwikkeling van DD in kinderen en volwassenen en naar de samenhang met andere factoren zoals intelligentie, executieve functies, en creatieve prestaties. Echter, het grootste deel van deze studies heeft DD onderzocht bij volwassenen of leerlingen vanaf 6-7 jaar. Er is weinig onderzoek gericht op DD bij jongere kinderen, hoewel de eerste tekenen van DD al bij tweejarigen werden gevonden. Ook is eerder onderzoek vooral gericht geweest op de uitkomsten, ofwel op de producten of ideeën die men genereert tijdens DD. In een dergelijke ‘product’ aanpak, worden de onderliggende denkprocessen van DD (grotendeels) genegeerd, waardoor we nog weinig weten over deze onderliggende processen.

Het doel van dit promotieonderzoek was juist om meer inzicht te krijgen in hoe DD bij jonge kinderen zich ontwikkelt en welke denkprocessen hierbij een rol spelen. Om dit doel te bereiken, is een groep van 107 kinderen gevolgd van begin groep 1 tot eind groep 2. Op vier momenten (met tussenpozen van zes maanden) is het DD van deze kinderen gemeten met behulp van de Alternatieve Toepassingen Taak (ATT), een van de meest gebruikte DD testen. Hierbij is niet alleen het product, maar ook het proces in kaart gebracht. Bij de laatste meting hebben de kinderen ook enkele tests gemaakt om hun executieve functies (EFs) te meten. Met de verzamelde data van dit longitudinale onderzoek, is antwoord gezocht op de volgende hoofdvragen :

- (1) Hoe komen kinderen op originele ideeën tijdens DD?
- (2) Hoe verloopt de ontwikkeling van DD bij jonge kinderen?

Om inzicht te krijgen in hoe kinderen op originele ideeën komen tijdens DD, is onderzocht welke denkprocessen ten grondslag liggen aan het genereren van originele ideeën bij kinderen (zie *Hoofdstuk 2*). Tijdens het uitvoeren van de ATT, werden de vierjarige kinderen in een interactieve dialoog aangemoedigd om zo veel mogelijk ongewone toepassingen voor alledaagse voorwerpen (bijv. een baksteen) te noemen en tevens aan te geven hoe ze op deze ideeën kwamen of wat ze daarbij dachten. Uit de analyses van de verbale uitingen kwam naar voren dat kinderen niet altijd konden uitleggen hoe ze op een bepaalde toepassing waren gekomen. Bij ongeveer de helft van alle toepassingen werd geen toelichting gegeven of wezen de uitingen niet op een specifiek denkproces. Voor de toepassingen die wel gerelateerd konden worden aan een denkproces, bleek dat de meeste werden gegenereerd op basis van automatische, associatieve processen die gebaseerd zijn op het ophalen van informatie uit het semantische en episodische langetermijngeheugen (dit proces wordt in het vervolg benoemd als *ophalen uit geheugen*). Daarnaast werden sommige

Samenvatting (summary in Dutch)

toepassingen gegenereerd op basis van meer complexe, top-down, executieve denkprocessen, zoals bijvoorbeeld het uitvoeren van mentale bewerkingen op een stimulus (bijv. de stimulus in delen uit elkaar halen of delen bij elkaar voegen of mentaal roteren; dit proces wordt verder *mentale bewerkingen* genoemd). De resultaten van de multiple en multilevel regressieanalyses lieten zien dat het aantal antwoorden dat kinderen gaven, door beide processen werd voorspeld, terwijl de originaliteit van de antwoorden alleen consequent door het proces ‘mentale bewerkingen’ werd voorspeld. Daarnaast werd gevonden dat vierjarige kinderen al een volgorde effect vertoonden in de mate van originaliteit van hun antwoorden, vergelijkbaar met volgorde effecten die in ander onderzoek bij oudere kinderen en volwassenen werden vastgesteld. Dit volgorde effect houdt in dat de originaliteit van de genoemde toepassingen toenam gedurende de reeks van gegenereerde toepassingen. Vergelijkbare volgorde effecten werden ook gevonden in de executieve DD processen, zoals mentale bewerkingen: hoe later in de reeks een bepaalde toepassing werd gegenereerd, hoe vaker de kinderen naar deze processen verwezen in hun uitleg over hoe ze op deze toepassing waren gekomen. Ook lijken de resultaten te wijzen op een omgekeerd volgorde-effect bij het denkproces ‘ophalen uit geheugen’: kinderen verwezen hier in het begin van de reeks veel naar, maar dit werd minder na de eerste paar gegeven toepassingen. Samen suggereren deze resultaten dat originaliteit afhankelijk is van een toenemend gebruik van executieve processen. Het executieve proces mentale bewerkingen, waarbij in gedachten eigenschappen of delen van voorwerpen worden geïsoleerd en vervolgens opnieuw gecombineerd, is daarbij een kernaspect in het bedenken van originele ideeën.

Als vervolg op de eerste studie is onderzocht hoe DD zich verder ontwikkelt bij vijf- tot zesjarigen en hoe DD is gerelateerd aan executief functioneren (zie *Hoofdstuk 3*). Dezelfde groep kinderen uit Hoofdstuk 2 werd anderhalf jaar later opnieuw getest met de ATT. Ook nu werden de genoemde antwoorden gecodeerd op originaliteit en werd nagegaan hoe vaak de kinderen hun antwoorden verklaarden uit de twee eerder gevonden denkprocessen: ophalen uit geheugen en mentale bewerkingen. Uit de multilevel regressieanalyse kwam opnieuw een volgorde effect in de originaliteit van toepassingen naar voren en ook hier bleek het proces mentale bewerkingen gerelateerd aan de originaliteit van de antwoorden. Daarnaast werd in deze studie ook een relatie gevonden tussen ophalen uit geheugen en originaliteit.

Naast het onderzoek naar de rol van de domein-specifieke denkprocessen tijdens DD, werd ook onderzocht hoe domein-algemeen executieve functies (EFs) gerelateerd zijn aan DD. Hiervoor voerden de kinderen een serie taken uit om diverse EFs te meten: werkgeheugen, shifting, inhibitie, en selectieve aandacht. Uit de analyses kwam naar voren dat geen van de gemeten EFs originaliteit voorspelde, behalve selectieve aandacht. Selectieve aandacht bleek een modererend effect te hebben op het volgorde-effect in originaliteit, maar anders dan verwacht. Kinderen met een zwakkere selectieve aandacht genereerden meer originele ideeën dan kinderen met een sterkere selectieve aandacht en dit effect was duidelijker in de late fase van het bedenken van ideeën bij een stimulus. Deze resultaten, samen met de bevindingen uit eerdere analyses, laten zien dat zowel domein-specifieke als

domein-algemene executieve processen een rol spelen in het ontstaan van originele ideeën tijdens het maken van de ATT. Het proces mentale bewerkingen is vooral positief gerelateerd aan de originaliteit van vroeg in de reeks gegenereerde ideeën, terwijl selectieve aandacht vooral negatief gerelateerd is aan de originaliteit van later in de reeks gegenereerde ideeën. Deze bevindingen lijken strijdig met de conclusies van *Hoofdstuk 2* waarbij oplopende originaliteit afhankelijk was van een toenemend gebruik van mentale bewerkingen. Mogelijk speelt hierbij een rol dat kinderen de ATT inmiddels vaker gedaan hadden en beter gewend waren aan de testomgeving en de taakeisen van ATT beter begrepen, waardoor ze wat nauwkeuriger de instructie volgden om echt nieuwe toepassingen te bedenken. Dit zou betekenen dat de kinderen op vijf- tot zesjarige leeftijd de taak anders hebben aangepakt dan toen ze vier jaar oud waren. Toekomstig onderzoek zal verder moeten uitwijzen hoe het proces van ideeën genereren bij kinderen verloopt en zich ontwikkelt over tijd.

In *Hoofdstuk 4* was de ontwikkeling van DD tussen vier en zes jaar onderwerp van onderzoek. De kinderen werden vier keer, iedere zes maanden, getest met de ATT. Opnieuw werden de productmaten vloeiendheid en originaliteit, en de procesmaten ophalen uit geheugen en mentale bewerkingen gecodeerd voor ieder kind. De ontwikkeling van DD werd onderzocht door de lineaire en kwadratische effecten van leeftijd op zowel product- als proces-gebaseerde metingen te toetsen met behulp van multivariate multilevel regressieanalyses. Uit de resultaten bleek dat kinderen zich in deze leeftijdsperiode sterk ontwikkelen in vloeiendheid en originaliteit. Ook namen de onderliggende denkprocessen toe met leeftijd: de kinderen gingen in hun toelichting steeds vaker verwijzen naar deze processen.

Naast het hoofddoel, inzicht krijgen in hoe kinderen op originele ideeën komen (micro-ontwikkeling) en hoe DD zich ontwikkelt bij jonge kinderen (macro-ontwikkeling), was een bijkomend doel van dit promotieonderzoek om de mogelijkheden tot het stimuleren van creativiteit bij jonge kinderen in het onderwijs te onderzoeken. In *Hoofdstuk 5* wordt een interventieprogramma gericht op creativiteit beschreven, namelijk het ‘Learn to Think programma for Preschoolers’ (LTT-P). Dit programma is ontwikkeld om creativiteit te bevorderen bij kleuters in China, juist omdat het Chinese onderwijs de kritiek krijgt dat het creativiteit belemmert – meer dan in andere landen. Omdat in China het voorschoolse kleuteronderwijs vaak georganiseerd wordt in drie niveaus (drie tot vierjarigen zitten in het junior niveau, vier tot vijfjarigen in het midden niveau, en vijf tot zesjarigen in het senior niveau), bestaat LTT-P uit twee delen met activiteiten voor kinderen van de midden en senior niveaus. Om de effectiviteit van LTT-P te evalueren, werd een quasi-experimentele studie uitgevoerd in een school in een stad in het noordwesten van China. In totaal hebben 68 kinderen uit het midden niveau en 87 kinderen uit het senior niveau deelgenomen aan deze studie. Het programma bestond uit een of twee activiteiten per week gedurende een periode van zes maanden. De ‘Lines’ en ‘Circles’ deelttests van de Torrance Test of Creative Thinking werden gebruikt om de vooruitgang in creativiteit tussen de voor- en nameting te meten. De antwoorden werden gescoord op vloeiendheid, originaliteit, en gedetailleerdheid. De

Samenvatting (summary in Dutch)

resultaten suggereren dat het LTT-P programma potentieel heeft om creatief denken bij jonge kinderen te bevorderen, voornamelijk met betrekking tot de dimensies originaliteit en gedetailleerdheid. De resultaten op vloeiendheid waren minder duidelijk. De positieve invloed van LTT-P op originaliteit lijkt veelbelovend voor het kleuteronderwijs in China.

In *Hoofdstuk 6* worden de bevindingen van de vier empirische studies samengevat en tevens worden de theoretische en praktische implicaties van deze bevindingen besproken. In relatie tot de vragen die zijn opgesteld in de Introductie (*Hoofdstuk 1*), hoe kinderen op originele ideeën komen, hoe DD zich ontwikkelt bij jonge kinderen, en hoe DD bevorderd kan worden, zijn er een aantal conclusies getrokken. Ten eerste blijkt dat kinderen van vier jaar oud al in staat zijn divergent te denken en dat vermogen groeit snel in de kleuterperiode. Door ze te stimuleren, kunnen kinderen op de leeftijd tussen vijf en zes jaar oud al een niveau van vloeiendheid van DD bereiken dat vergelijkbaar is met DD bij scholieren en volwassenen. Daarnaast kunnen gestructureerde interventie programma's als LTT-P een waardevolle bijdrage leveren aan het onderwijs van DD en creativiteit in de praktijk. Verder blijken verschillende typen denkprocessen betrokken bij DD waarin de processen ophalen uit geheugen en mentale bewerkingen de belangrijkste zijn voor het genereren van originele ideeën. Er werd het meest naar het proces ophalen uit geheugen verwezen toen kinderen uitlegden hoe ze op hun ideeën waren gekomen. Daarnaast was het proces mentale bewerkingen, hoewel hier veel minder vaak naar verwezen werd door kinderen, positief gerelateerd aan het genereren van originele ideeën in DD. Ook lijkt het er op dat het genereren van originele ideeën verbonden is met een relatief diffuse aandacht voornamelijk in de latere fase van idee genereren. Als laatste worden in *Hoofdstuk 6* beperkingen van dit promotieonderzoek en suggesties voor mogelijk vervolgonderzoek toegelicht. Uiteindelijk laten de resultaten beschreven in dit proefschrift zien hoe we beter grip kunnen krijgen op de aard van DD, als kernaspect van creativiteit, door de onderliggende denkprocessen te analyseren. Toekomstig onderzoek is nodig om hier beter zicht op te krijgen, waarbij aanbevolen wordt om met behulp van andere onderzoekparadigma's en het liefst in realistische, ecologisch valide omstandigheden dit thema verder te onderzoeken.

中文摘要

(Summary in Chinese)

发散性思维通常被定义为可以支持人们从不同的方向及角度对事物进行探索和思考的一种思维过程。作为创造力的重要组成部分，发散性思维对原创或新颖性想法的产生尤为重要，并被广泛认为是创新和问题解决的重要贡献因素。由此，发散性思维在近几十年受到了研究者的高度关注。发散思维测验被开发并广泛地应用于测量个体差异。同时，采用这些测验，以往的研究也关注于测量及描述发散性思维在不同年龄群体或跨年龄群体的发展水平，揭示了发散性思维与流体（及言语）智力、认知控制功能以及实际生活中的创造力成就等之间的相关及协变模式。然而，现有关于发散性思维的研究仍局限于两个方面。首先，目前的研究主要侧重于成人及学龄儿童而鲜少专注于学前儿童，即便有研究表明一到两岁的幼儿已经能够进行发散性思维。其次，现有研究大多从产品的视角来对发散性思维进行研究，关注于被试在发散思维测验中所构建的有形、实际的或停留在思考层面的产品，并基于此对其发散性思维从流畅性、灵活性以及独创性等维度进行评价。基于这一产品视角，被试在进行发散思维时的具体思考过程往往被忽视，也因此尚未能被很好的理解。从这两个局限出发，本论文的首要目标是从过程的视角对学前幼儿的发散性思维进行研究。为此，我们对 107 名幼儿进行了为期一年半的追踪研究，从四起到到六岁为止。在这期间，我们采用了国际上广泛使用的发散性思维测验，即多用途测验（Alternative Uses Task），对幼儿进行了四次测验，每半年一次。在最后一次测验中，我们同时采用了一系列标准化测验对幼儿的认知控制功能进行了测验。基于这一追踪研究所采集的数据，我们旨在探讨两个问题：

- 1、从微观层面出发，儿童在发散思维中如何产生原创性想法？
- 2、从宏观层面出发，发散性思维在学前阶段如何发展？

为研究儿童在发散思维中如何产生原创性想法，我们首先分析了儿童在想法生成过程中的思维过程。本论文的第二章对该研究进行了详细阐述。我们采用了多用途测验，要求四岁幼儿对若干日常物品（例如砖头）给出尽可能多的、不常见的用途。此外，在给出这些用途的同时，我们采用了互动性对话来鼓励幼儿对他们的思维过程进行报告。基于对其话语的内容分析，我们发现幼儿并不总是能够很好地解释他们是如何想到一个特定的用途。这反映在，对于大约一半的用途，幼儿并没有给出解释或者其给出的解释不够清楚而难以对其编码。而对于那些有明确解释的，大部分用途是幼儿基于自动地、自下而上的联想性思维过程（associative processes）而给出的，涉及对语义或情景长时记忆的提取（以下简称“记忆检索”过程）。除此以外，偶尔也有一小部分用途是幼儿基于若干类费力的、自上而下的控制性思维过程（executive processes），例如对刺激物品进行分解、再组合及旋转等心理操作（以下简称“心理操作”过程）。采用多维回归及多水平逻辑回归分析，我们进一步发现，首先，无论是联想性还是控制性过程都预测了儿童在多用途测验中的流畅性分数，然而只有心理操作这一控制性过程预测了独创性。其次，四岁幼儿所给出用途的独创

性已经呈现出一个序列效应 (serial order effect)，即在幼儿针对一个物品给出的一系列用途中，越是在这个系列的后面，则一个用途的独创性越高。最后，类似的序列效应也呈现在各类控制性过程的应用之中。对于一个特定的用途，其越晚出现在用途系列之中，则幼儿越有可能解释说他们是基于控制性过程而想到这个用途。特别的是，我们观察到独创性的序列效应和心理操作这一控制性过程的序列效应呈现出非常相似的时间模式。与此相反，记忆检索过程的发生则呈现出一个反向但并不显著的序列效应，即针对一个特定的刺激物品，幼儿指向该过程来解释所给出用途的概率在初始阶段较高，但在给出最初几个用途之后，这一概率反而随着上升的用途排序而下降。综上，这些研究结果显示，持续上升的用途独创性很有可能依赖于可控制性过程在想法产生过程中愈加频繁的参与。尤为重要的是，心理操作这一控制性过程，涉及（1）对物品属性或组成部分进行分离以及（2）之后对这些属性及组成部分进行再组合以产生一个新的、结构化的物品的，很有可能是发散性思维一个重要特征并尤其支持其中原创性想法的产生。

作为后续研究，我们在第二个研究中更加深入地检验了原创性想法在发散思维中实时展开的过程以及控制性过程在其中所起的作用。本论文的第三章对该研究进行了详细的阐述。本研究与上一研究的被试为同一组幼儿，只是在本研究中，幼儿的平均年龄比其在上一研究中大约大了一岁半，居于五到六岁之间。同样地，我们采用多用途测验对这些儿童的发散性思维进行测验，要求他们尽可能多地针对日常物品给出不常见的用途并报告自己的思维过程。随后，对每一个给出的用途，我们对其独创性进行了评分。同时，根据幼儿的解释，针对每一个用途，我们还对其产生过程是否与记忆提取及心理操作这两个在上一研究中界定的重要发散思维过程相关分别进行了编码。通过多水平逻辑回归分析，我们再一次发现幼儿给出用途的独创性存在序列效应。同时，心理操作这一过程再一次显著地预测了用途的独创性。这些研究结果重复了我们的上一研究。然而，除了心理操作以外，在本研究中我们发现记忆提取过程对用途的独创性也具有正效应。

除了对记忆提取和心理操作这两个领域相关的过程是如何在发散性思维中起作用进行研究，本研究的另一个重要目标是考察领域一般的认知控制功能 (executive functions) 对原创性想法实时展开过程的影响，以及这些认知控制功能与记忆提取和心理操作这两个发散性思维过程之间的相互作用。为此，我们采用了六个标准测验对幼儿的认知控制功能包括工作记忆、认知转换功能、认知抑制功能及选择性注意进行了测量。研究结果显示，在这些认知控制功能之中，没有任何一个对独创性有预测效应。但是，选择性注意对独创性的序列效应存在调节作用，体现为选择性注意水平较低的儿童往往能够给出更多或更具有原创性的想法，而且这一效应在想法生成的后期变得愈加明显。综合这些以及上面已提及的研究结果，我们总结得出，不论是领域一般的认知控制功能还是领域相关的发散思维过程都对发散思维过程中

原创性想法的实时展开起作用。其中，控制性思维过程（尤其是心理操作）的早期介入可能与五到六岁幼儿相较于四岁时愈加遵从测验任务要求——即给出新的、原创性的想法——相关，而在想法生成的后期，更加宽泛的注意范围（即相对不集中的注意力）则可能代替控制性思维过程，可能是因为后者逐渐枯竭，来促进原创性想法的生成。该结论与我们基于上一研究（第二章）所得出的设想并不一致。在之前针对四岁幼儿的研究中，我们曾总结出持续上升的独创性主要依赖于控制性思维过程愈加频繁地介入于发散思维之中。我们推断，这些不一致的研究发现可能体现了幼儿发散思维的发展性变化。在多次参与多用途测验之后，幼儿可能变得更加熟悉测验环境并能更好地理解多用途测验要求他们给出真正新颖的用途的要求。由此，五到六岁的幼儿可能采取了不同于他们在四岁的时候所使用的方式来完成多用途测验。未来研究应更细致的检验想法生成的具体思维过程在儿童中是如何变化和发展的。

与前两个研究侧重于发散性思维的微观发展不同，在第三个研究之中我们关注了发散性思维在学前阶段的宏观发展。本论文的第四章对此进行了详细的阐述。在该研究中，我们使用了在前述追踪研究总所采集的发散性思维数据。具体而言，我们采用多用途测验对幼儿的发散性思维进行了四次测验，每半年一次。如前所述，我们要求幼儿对日常物品给出尽可能多不常见的用途并同时报告他们的思维过程。基于幼儿给出的用途，我们从产品的角度对幼儿发散性思维的流畅性和独创性进行了评分。除此以外，基于幼儿对其思维过程的解释，我们同时从过程的角度对幼儿使用记忆提取和心理操作两个发散思维过程进行了评分。随后，我们采用了多元多水平回归分析对幼儿发散性思维的发展轨迹进行了分析，分别检验了幼儿年龄对于基于产品和基于过程的发散性思维测量分数的线性及二次方效应。我们的结果表明，随着年龄的增长（四到六岁），幼儿在流畅性和独创性上有非常明显的发展。同时，随着年龄的增长，幼儿亦更加频繁地提及记忆提取及心理操作这两个发散思维过程来解释他们是如何想到给出的用途的。

综上所述，本论文的前三个研究专注于对幼儿发散性思维的思维过程进行研究，以更好地理解幼儿是如何产生原创性想法以及发散性思维在学前阶段如何发展。除此以外，本论文亦旨在探讨如何在学校教育中培养幼儿的创造力。由此，在本论文的第五章中，我们介绍了一个结构化的创造力学前幼儿培养项目，即学思维幼儿培养项目（Learn to Think program for Preschoolers）。鉴于常常有批评认为中国的教育体系相较于其他国家的教育体系而言更倾向于阻碍学生的创造力发展，该项目专为中国幼儿园量身制作，旨在为幼儿园实施创造力教育提供支持。在中国，学前教育通常由三个年级构成。其中，三到四岁幼儿上小班，四到五岁幼儿上中班，而五到六岁幼儿上大班。学思维幼儿培养项目设计了对应于中班及大班幼儿的系列活动。为了检验该项目对创造力培养的成效，我们开展了一个准教学实验研究并结合采用了

前后测的研究设计。该教学实验开展于中国西北部的一个城市幼儿园，涉及 68 名中班幼儿以及 87 名大班幼儿，为期两个学期。在这期间，根据幼儿园的安排，每周每年级开展一到两个学思维活动。在实验的前后测中，我们分别采用了托兰斯测验（Torrance Test of Creative Thinking）图画版中的平行的线条测验和圆形测验来对幼儿创造力进行测量，得到了流畅性、独创性和精进性三个维度分数。我们采用了重复测量多元多因素方差分析（repeated measures multivariate analysis of variance）分别对学思维项目对中班及大班幼儿创造力发展的影响进行了检验。分析结果表明，学思维幼儿项目对幼儿创造力的发展尤其在独创性和精进性维度上有潜在的积极的作用。然而，本研究中和流畅性相关的结果则不甚清晰，难以明确该项目是否可以促进幼儿在流畅性上的发展。我们认为，学思维幼儿项目对幼儿独创性发展的积极作用对促进中国当前的学前创造力教育尤为相关。

最后，在本论文的最后一章即第六章中，我们对前述四个实证研究的结果进行了总结并讨论了其在理论及应用上的意义。儿童如何产生原创性想法？发散思维在学前阶段如何发展？此外，我们该如何培养儿童的发散思维，或者更宽泛地说，如何培养儿童的创造力？基于本论文中各个研究的结果，我们可以给出若干主要结论。首先，从宏观发展的角度而言，四岁幼儿已经能够很好地进行发散思维，而且这一能力在学前阶段得到迅速发展。尤其是当我们为幼儿提供恰当的鼓励和刺激时，五到六岁的幼儿的发散思维，反映在他们的平均流畅性分数上，已经达到可以和学龄儿童或成人相较的水平。同时，对于创造力培养的教育实践而言，向幼儿园提供类似于学思维幼儿项目等结构化的支持不失为一种有效且可行的模式。其次，从微观发展的角度而言，发散性思维涉及了不同类别的思维过程，而这些过程中又尤属记忆提取和心理操作两个思维过程最为重要。其中，记忆提取最经常被幼儿提及，解释说他们是基于这一过程而给出特定的用途。反观心理操作过程，虽然很少被提及，却被发现尤其与原创性想法的产生息息相关。最后，除了心理操作这一费力的控制性发散思维过程，原创性想法的产生，尤其在其后期阶段，与个体更为宽泛的注意范围也紧密相关。最后，在第六章中，我们对本论文所囊括研究中最主要的局限与不足进行讨论，并对后续及未来研究给出建议。本论文的研究展示了如何从过程的角度来研究发散思维，以加深我们对发散思维以及创造力本质的理解。我们鼓励未来研究更多地从类似的过程的角度并结合其他研究任务与范式来研究创造力，最好能够同时结合现实的生活情境以提高研究的生态效度。

Acknowledgement

Although I officially started this PhD project in September 2015, I had begun with the preparation of the research proposal about a year earlier. The very initial idea of this proposal came from a discussion I had with Xinfu Yi, a caring teacher and a passionate creativity researcher. We thought that the traditional field experimental design provides too little information regarding how teaching interventions take effect and thus, speculated that it would be wonderful to use ‘micro-experiments’ in classrooms, which should tell us whether and how a class is effective. Since then the idea of observing the process (of teaching and learning) was planted in my mind. The more concrete shaping of this initial idea took place when I learned about the micro-genetic study from a guest lecture given by Ziqiang Xin and the think-aloud method from the article of Gilhooly and colleagues (2007). I sincerely thank these researchers for the inspirations they have brought me. Finally, these ideas were developed into a proposal. Here I would like to particularly thank John Xuexin Zhang, a pioneering higher education reformer and a wonderful teacher. Although I was not your student, you have extensively and critically reviewed my proposal and spent time to help me polish it. You were also always generous in giving me advice and encouragement during my PhD. If I am to be a teacher in the future, I would like to be your type!

With the proposal I started to apply for universities. At then, I met Henk van Rinsum and Marcel van Aken at the PhD orientation in Beijing. Thank you Henk and Marcel! It was you who brought my application to Utrecht University (UU) and made a bond between UU and me. Here I would also like to acknowledge the China Scholarship Council (CSC) and the UU-CSC program. Without the scholarship provided by CSC and the support from the UU-CSC collaboration, this PhD project would not have taken place. Especially by witnessing the recent increase of doubts and negative discourses on the media about Chinese scholars abroad, the value of such collaborative programs became increasingly salient to me. I would like to thank and express my deepest respect to all involved parties that have built and kept working on enhancing such partnerships.

For the completion of this PhD project I own a great deal of gratitude to my supervisors Evelyn Kroesbergen, Paul Leseman, and Hanna Mulder. Evelyn, you were the first one in contact with me and agreed to supervise me. With your support I was successfully granted with the CSC scholarship so that this project could start. Being literally your first foreign PhD student, I truly appreciate that you have invited me to come regardless of potential cultural shocks, taken efforts to make me feel included in this new environment, and continuously encouraged me and reminded me of my strengths when I felt down. Paul, we met when I arrived at the Netherlands. As Evelyn moved to Radboud University half way and thus became less involved in my project, you are the one who was there from the beginning until this last point. As critical, knowledgeable, and imaginative and creative as you are, working with you has been a real adventure for me. Excited in learning new things while sometimes also feeling upset of not being able to go a step further, I have had a complete, 360-degree, 3-D or n-D experience of doing a PhD. In addition to all the professional guidance you have offered me in the last years, I am particularly grateful that,

Acknowledgement

at a few times when I faced great challenges, you were there to listen to me and to help me. Hanna, I first met you at the dust-to-gold meetings. I could not imagine then that you would become my supervisor a while after that. You are a wonderful supervisor, responsive, caring, modest... You have been so involved: you brought fresh inputs which not only broadened my view but also reshaped and boosted this project, you worked closely with me to sort out all those small, tiny details, and oftentimes, you also took me hand in hand to practice essential writing skills. Without you joining the team, the dissertation would have looked very different. Dear Evelyn, Paul, and Hanna, you have made a wonderful supervisor team with great diversities: Evelyn distinguished by a romantic approach, Hanna featured with rigorous thoroughness, and Paul maximized with critical questioning. No wonder that the diversities have sometimes posed challenges for us to agree with each other. Nonetheless, there is no way to deny that these diversities have also pushed us to practice the topic of this PhD project—thinking outside of the box—and made this project much more interesting! With all my heart, thank you!

Apart from my supervisors, another source of academic inspirations and daily comforting came from my colleagues of the creativity group and my officemates. Dear Eveline, Marije, Marloes, and Mare, we met each other due to our common interests in creativity research, but the bond (be it collegueship or friendship) between us soon grew beyond that. As a small community, we helped each other to grow professionally, organized conferences together, and supported each other whenever or whoever needed this. We have also spent time together by as daily as having a coffee at a nice café or as special as paying a baby visit (“kraambezoek”) to one another. And the ‘sleeping-trip’ I had with you in America is absolutely unforgettable. As an initially rootless PhD here, being a member of the creativity group has truly created a feeling of belonging for me, making me feel supported and connected. Dear Spark, Lijie, Asli, Isabelle, and Ruben, we have shared the office of E2.06 in different periods. There are quite some diversities among us: different backgrounds, different research interests, and different occupations. However, these diversities only made our conversations much more fun and dynamic. I enjoyed all the small things we have done together: having lunch, complaining about the Dutch weather, talking in my broken Dutch,..., and last but not least, growing baby plants in the office. Thank you all!

There are also many others I would like to thank. Mirjam Moerbeek and Donald Hedeker: thanks to your support in statistics, my PhD progress has been significantly accelerated. Students and research assistants who have participated in data collection: without your efforts and hard work, this project would not have been finished. Moreover, my greatest appreciation should be given to the children, parents, and schools who have participated in this project. Specifically, I would like to thank Jan Willem Helmink, Noortje Verduijn, Marlies Theunis, Hanny de Kort, Angela Schiranian, Gerion Goossens, Alexandra van Dijk, and Derk Lettink for the support you have so generously provided me in recruiting the participants and in arranging practical matters for the data collection. Furthermore, I should certainly not forget to mention the extensive support I have received from other colleagues

at the department. To name a few, Eric and Jos, thank you for all the technical support; Natalie, thank you for arranging all the administrative matters; Margriet, thank you for all the small, heart-warming talks we had in the hallway; Ora, thanks to joining your Intake project, I got new research inspirations...Moreover, fellow PhDs, I enjoyed very much the lunch time with you; and the social group, I liked a lot of the nice social events you organized! I would also like to thank fellow Chinese PhDs, Xiaoyan, Mei, Xiaojing (Gu), and Xiaojing (Sun). It was wonderful to meet you and to get to know you here in the Netherlands.

While completing this PhD, a number of major changes have taken place in my life. I own a great deal of gratitude to friends and families who were always there for me, at good and bad times. Friends I have made here outside of the university, Hein, Eirini, Bernard, you have become a major part of my private life. Hein, from discussions with you, I learned a great deal of cultural differences and the world history. Eirini, your Greek may-be-time life style has truly taught me to look at things differently over the last years. Now with also Nikos and the little Jorgos, visiting you has become an event for me and Armel, creating always cheerful memories. Bernard, you are a 'crazy' person but surely a great friend. Although we always argue once we see each other, the feeling of no distance really makes you more a family than a friend to me and Armel. Also old friends from China, A-meng, Haishan, Linming, and Xiaofang, even from a distance, I have always felt supported by you.

I owe many thanks to my parents. Doing a PhD was more a desire of mine than yours. Being in the different time zones, holding different goals and habits, having different life styles...the enlarging differences have inevitably introduced some conflicts between us. I am grateful that, regardless of the conflicts, you have always tried to be supportive and gave me the space to be myself and to do my things. I am also immensely grateful to my dear sister and brother. Although you have very little ideas about what I am actually doing and may sometimes even wonder why I would do this PhD, you have always been supportive in a variety of manners. My gratitude should certainly be extended to my parents-in-law. Merci Martine et Alain! With all the great food and the sweet conversations, you have so generously offered me a second home here. And I certainly enormously enjoyed every funny joke you made or you inspired me to make over the stressful PhD years. At last I want to thank my husband. Dear Armel, it was simply the best to have met and married you in this PhD and in my life.

About the Author

Honghong Bai was born on September 11th, 1988, in Pingyang, Zhejiang, China. She obtained her secondary school degree from Pingyang High School in 2007. In the same year, she moved to Guangzhou and started her study in Computer Science and Technology at Southern China Normal University. Due to growing interests in how people's mind functions, she started to study psychology by herself in the university. In 2011 she obtained her Bachelor's degree and started to work in Zengcheng College of Southern China Normal University (now Guangzhou College of Commerce) as a network manager and IT supporter. One year later, she left her job and moved to Xi'an to start her Master's study in Developmental and Educational Psychology at Shaanxi Normal University. In this period, her broad interests into people's mind was narrowed down and led her to eventually focus on *creativity in young children*. Within this specific topic, she developed her PhD proposal on the understanding and the development of creative thinking in young children. In 2015 she obtained her Master's degree and received a scholarship from the China Scholarship Council. On the 1st September 2015, she arrived at the Netherlands and started her PhD research in collaboration with Evelyn Kroesbergen, Paul Leseman, and Hanna Mulder at the Department of Education and Pedagogy, Special Education: Cognitive and Motor Development (now Development and Education of Youth in Diverse Societies) of Utrecht University. The research findings have been presented at conferences and (submitted to be) published in international journals. During the PhD period, Honghong has worked as a part-time lecturer at Utrecht University of Applied Sciences (Hogeschool Utrecht; August 2019–August 2020) and as a part-time researcher for the Intake project within the department (June–December 2019 & September–December 2020). She will continue pursuing her research interests in creativity in her future career.

List of Publications Related to This Thesis

Scholarly Publications

Bai, H., Mulder, H., Moerbeek, M., Kroesbergen, E. H., & Leseman, P. P. M. (2021). Divergent thinking in four-year-old children: An analysis of thinking processes in performing the Alternative Uses Task. *Thinking Skills and Creativity*, 40, 100814. <https://doi.org/10.1016/j.tsc.2021.100814>

Bai, H., Leseman, P. P. M., Moerbeek, M., Kroesbergen, E. H., & Mulder, H. (2021). Serial order effect in divergent thinking in five- to six-year-olds: Individual differences as related to executive functions. *Journal of Intelligence*, 9(2), 20. <https://doi.org/10.3390/jintelligence9020020>

Bai, H., Duan, H., Kroesbergen, E. H., Leseman, P. P. M., & Hu, W. (2019). The benefits of the Learn to Think program for preschoolers' creativity: An explorative study. *Journal of Creative Behavior*, 3, 699–710, <https://doi.org/10.1002/jocb.404>

Professional Publications

Hu, W., Shan, X., **Bai, H.**, & Duan, H. (2016). 学思维幼儿活动课程 教师用书[Learn to Think activity curriculum for young children (Teacher's book)]. Beijing: Foreign Language Teaching and Research Press