# The Development of Preschoolers' Semantic Strategy Use on a Divergent Thinking Task

Final version

Master's thesis

Utrecht University

Course: Master's programme in Clinical Child, Family and Education Studies

Student: A. W. Oppeneer

Student number: 5742986

Supervising lecturer: H. Bai

Second assessor: I. Friso-van den Bos

Date: 08-06-2018

Abstract

**Aim.** To date there is still little known about the development of preschoolers in terms of their strategy use in divergent thinking (DT). Therefore, the aim of this study was to investigate the longitudinal relationship between semantic strategy use and the quality of DT in preschool children, including the effect of gender. **Method.** This relationship is investigated through the results of the Alternative Uses test (AU) of 35 children (14 boys), 3.94-4.88 years old. These children were tested twice within approximately five months. **Results.** Semantic strategy use appeared to be only a significant predictor of DT quality at the first measurement, not at the second. In addition, children did not use the semantic strategy more often over time. The quality of DT on the other hand did increase significantly between the first and second measurement (partial $\eta^2 = 0.288$). Furthermore, there were no gender differences for both the semantic strategy use and the quality of DT. Finally, the quality of DT at the second measurement could not be predicted by how often the participants used semantic strategy at the first measurement. **Conclusion.** Altogether, it can be concluded that the quality of DT and semantic strategy use develop differently, resulting in no clear longitudinal relationship between DT quality and semantic strategy use.

*Keywords:* divergent thinking, semantic strategy use, longitudinal, gender, alternative uses test.

# The Development of Preschoolers' Semantic Strategy Use on a Divergent Thinking Task

Creativity is becoming increasingly important in the rapidly changing society nowadays (Binkley, et al., 2012). Employees are more and more expected to be able to gather information, compare it, and to come up with innovative ideas. In preparation for this, primary school children are taught the so-called 21st century skills. These skills include, in addition to ways of working, ways of thinking: creative thinking, problem solving, and metacognition (Larson & Miller, 2011). An increasing number of schools see the value of 21st century skills and is teaching them to their students.

## Creativity and Divergent Thinking

Even though creativity is a hot training topic in educational practice, how creativity is operationalized and measured is still an issue in creativity research (Reiter-Palmon & Arreola, 2015), mainly because the concept 'creativity' is complex (Kim, 2011). According to the standard definition, creativity requires originality and effectiveness (Runco & Jaeger, 2012). However, this doesn't entirely cover the load. In fact, creativity can be useful in different domains, asking for different skills (Han, 2003). In line with that, Han (2003) reported, based on their study, that it is hard to predict someone's creative ability in one domain, based on his or her creative abilities in other domains. This implies that creativity might be domain-specific, which makes it difficult to research creativity as a whole. Nonetheless, creativity has been of main interest in many research projects in the last decades, using divergent thinking (DT) tasks for measuring creative thinking (Zeng, Proctor, & Salvendy, 2011). DT is considered, even though not the same, but as an important aspect of creative thinking (Russo, 2004). As a matter of fact, research suggests that these DT tests can provide useful estimates of humans' creative thinking potential (Dumas & Dunbar, 2014; Runco & Acar, 2012). Given the complexity of creativity, and the importance of DT, this study will focus on the aspect of DT.

## Divergent Thinking and Strategy Use in Early Childhood

DT is defined as the ability to generate numerous and diverse ideas to open-ended questions (Kuhn & Holling, 2009). Therefore, people have to make unexpected combinations and identify connections among remote associates (Benedek, et al., 2014a). Although DT has been widely investigated since 1970, there is a growing interest only in recent years in looking into the underlying cognitive processes, including strategy use.

A strategy can be defined as any procedure that is nonobligatory and goal directed (Siegler & Jenkins, 2014). Empirical studies (Beaty, Silvia, Nusbaum, Jauk, & Benedek,

2014; Gilhooly, Fioratou, Anthony, & Wynn, 2007; Nusbaum & Silvia, 2011) have suggested a strong connection between the strategies that adults use in DT and the originality of their ideas. Gilhooly and colleagues (2007), for example, reported that initial responses are based on a strategy of retrieval from long-term memory, but are therefore mostly not original. Later responses, which are more often novel and original uses, seem to be based on a small number of other strategies. There is no such research on children, to our best knowledge.

However, even though not exactly the same, we can take hints from the research on mathematical problem solving in (early) childhood. Based on a rich experience of researching thinking strategy in mathematical problem solving, Siegler and Jenkins (2014) concluded that children always, consciously or unconsciously, use strategies while solving a problem. This strategy use in problem solving requires the children to use their knowledge in a flexible and creative manner (Siegler, 2007), which underlies the production of original ideas (Benedek, et al., 2014a). Although mathematical problem solving is not the same as generating original ideas in DT, Russo (2004) suggests that DT requires a flexible use of knowledge as well. However, our knowledge about strategy use in DT is based on very limited data acquired with adults. Little is known about crucial periods and changes in strategy use, neither in mathematical, nor in non-mathematical problem solving, of children in the developmental age. According to the overlapping-waves model of Siegler (1996), concerning strategy use in the broad sense, the strategies that children use are slightly changing over time: When children discover new, more advanced strategies, they will slowly replace the older, less mature, strategies. Nonetheless, the generalizability of these findings is not clear. The aim of this research is consequently to gain more insight in the development of strategy use in the DT of young children.

**Experimental Strategy and Semantic Strategy**

Concerning strategy use in DT and the process of generating original ideas and creative solutions, Gilhooly and others (2007) investigated which strategies people use to come up with these ideas. For this purpose, they used the Alternative Uses test, a typical DT task, which requires participants to produce as many unusual uses for common objects (e.g., a spoon or a brick) as they could (Gilhooly et al., 2007). This can lead to ideas which are derived from the participants memory, and are therefore familiar to them, or ideas which are newly created during the task and are thus unknown for the participants themselves (Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014b). Strategy use underlies the generation of both common ideas and creative ideas (Gilhooly et al., 2007). To be able to identify which specific strategies participants use in the process of idea generation, some participants of Gilhooly and

others' (2007) study were placed in a think aloud group. All their answers were coded and grouped into four larger categories of strategies, which are described in more detail in Appendix I: Memory use production, property use production, broad use-based production, and disassembly use production.

These strategies can be divided into two main strategies: the experimental strategy (where the memory of a specific experience is invoked) and the semantic strategy (in which abstract conceptual characteristics are used to generate examples). The experimental strategy includes the memory use production, while the semantic strategy underlies property use, broad use-based, and disassembly use production (Gilhooly, et al., 2007). Research showed that there were less answers that were generated based on the semantic strategy and that they appeared later in the response sequence as well (Beaty & Silvia, 2012; Heinonen, et al., 2016). This can be explained by the fact that memory responses tend to occur automatically and rapidly as a response to the shown object (Hass, 2017).

However, all these results are based on the performances of university students. It is not clear how young children develop in using strategies in DT. Because the strategy use of young children is still developing into more advanced strategies (Alexander, Graham, & Harris, 1998), it is expected that preschoolers will use the semantic strategy more often over time. Memory responses are namely occurring automatically and are therefore less mature (Gilhooly, et al., 2007).

**Predictors of the Quality of Divergent Thinking**

One of the predictors often associated with the quality of DT is the age of the child, although researchers have not reached consensus on how DT quality is related to age. Some researchers (e.g., Bijvoet-Van den Berg & Hoicka, 2014) found a positive correlation, others found, in contrast, a decrease of DT quality over time. This last group (e.g., Gardner, 1982) proposed that preschool children have high levels DT abilities, but that these abilities decline when the children learn conformity after entering school. Other researches demonstrated the same decline, but also showed a subsequent increase after the fourth-grade (Claxton, Pannells, & Rhoads, 2005). Since theses researches have an overall initial increase of DT quality in common, it is expected that the quality of DT will increase during preschool.

Another predictor, which has been the subject of an increasing number of research, is the possible gender difference regarding the generation of creative solutions. Over the years, several studies (e.g., Kogan, 1974; Pilar Matud, Rodríguez, & Grande, 2007; Kershner & Ledger, 1985, have been carried out on the gender differences in DT. Although the research results are not always consistent, a growing body of recent literature indicates that females

have advantage in DT (Baer & Kaufman, 2008; Kuhn & Holling, 2009), particularly on tasks that necessitate verbal fluency (Abraham, 2016) or originality (Shi, Xu, Chen, & Qiu, 2017). Recent neuroimaging studies showed that males and females have different brain activity while performing a DT task (Abraham, Thybusch, Pieritz, & Hermann, 2014; Shi, et al., 2017), even though men and women had similar scores on that task (Razumnikova, Volf, & Tarasova, 2009). This indicates that males and females may differ from each other in using strategies while finishing DT tasks. This will hence be another interest that this study will explore. Since females seem to perform better on DT tasks, it is expected that they use semantic strategy more often than their male peers.

**The Present Study and Hypotheses**

There is still little known about the development of pre-schoolers in terms of their strategy use in DT, although it is assumed that semantic strategy use plays an important role (Heinonen, et al., 2016; Wang, et al., 2017). For this reason, the present study will investigate the longitudinal relationship between semantic strategy use and the quality of DT of preschool children.

Before answering this main question, we will first examine the concurrent relationship between semantic strategy use and the quality of DT in preschool children. This allows us to see whether we can replicate the results from adult studies (e.g., Gilhooly, et al., 2007) to a study with young children. Consequently, we expect that semantic strategy use positively influences the quality of DT in young children as well.

Secondly, we will investigate how semantic strategy use develops in young children and whether gender influences this development. Since gender seems to play a role in creative thinking (see for example Baer & Kaufman, 2008, and Kuhn & Holling, 2009), we suspect it might also play a role in the children's semantic strategy use. We expect an increase of semantic strategy use over time, with girls using the semantic strategy more often.

At last, based on the results of the first two analyses, we want to examine whether there is a longitudinal influence of the use of semantic strategy on the quality of DT while taking gender into consideration. An increase of DT quality is expected, with girls scoring higher on DT quality than boys. Due to the limited research regarding the longitudinal relationship of semantic strategy use and the quality of DT, it is impossible to formulate a strong hypothesis.

## Methods

**Participants**

To recruit participants, a letter that introduces the research has been sent to more than 20 schools in the Netherlands. Four schools eventually agreed to participate in the research. Parent consent was then sent via the schools. In the end, 107 children from four primary schools in the Netherlands participated in this study with parents' approvals. Among them, 57 children (24 boys, 3.85 – 4.79 years old, $M$ = 4.31, $SD$ = .27) were repeatedly measured (first in October or November 2016, second in March or April 2017). This study will hence use the longitudinal data of this part of participants.

**Measurements**

This study utilized the Alternative Uses test (AU) to evaluate the quality of DT and the thinking strategies that are used in the process of DT. The AU, developed by Guilford in 1967, is a psychometric measure which requires participants to generate as many original uses as possible for familiar objects. Although the AU is frequently used in creativity research, little is known about validity and reliability. The study of Jung and Ryman (2013) mentions an interrater reliability of originality measures ranging between 0.62 and 0.95. For validity, statistics of 0.51 and 0.52 have been reported on the factor of 'spontaneous flexibility' for adult samples (Domino & Domino, 2006). However, the extended literature using the AU, supports the assumption that this test is seen as a valid instrument for measuring DT.

In the current study, two sets of six pictures of daily objects are used. They were presented to the participants in a random order. The first set, used at the first measurement, consists of a toothbrush, a shovel, an umbrella, a car tire, a bin, and a pencil. In the second set, used at the second measurement, a wash cloth, a brick, a broom, a basket, a net, and a spoon are included.

**Test procedure.** The AU was administered individually and orally by trained test administrators. The testing takes 12-24 minutes per child, and is recorded with a camera. At both measurements, the test took place in a separate room at school. After an example, the children were asked to produce as many uses for six other items as they could think of. During the task, the children were asked how they came up with certain ideas.

**Assessing the quality of DT.** After scoring, the AU gives information on four levels of DT: Fluency (the number of given solutions), flexibility (the diversity of the solutions), originality (the novelty or statistical infrequency), and elaboration (when someone follows an associative pathway) (Kim, 2006; Runco & Acar, 2012). This study focusses on the quality of DT, which will be determined by the fluency and the originality scores.

**Coding process of DT strategies.** Before coding the longitudinal data, an adapted coding scheme based on Gilhooly and others (2007) for analyzing DT strategies in young

children has been developed via a pilot. The first author was afterwards trained to use this adapted coding scheme. Therefore, seven transcriptions of the pilot study were segmented into separate episodes and coded by both the first author and the researcher of the pilot study. As this study concerns difference between the experimental strategy and the semantic strategy, the five last-mentioned general categories, see Appendix I, are combined and further generalized to the semantic strategy. Memory based production represents the experimental strategy.

After the training to use the adapted coding scheme, the interrater reliability was determined based on the basic categories as well as the grouped categories. Cohen's kappa showed not only a substantial agreement between the two researchers when comparing the basic categories, $\kappa = .767$, $p < .001$, but also when viewing the grouped categories, $\kappa = .832$, $p < .001$.

To be able to include the use of semantic strategy as a separate factor in this study, we created a new score (in percentage). This percentage displays, for each participant, the use of semantic strategy relative to the total frequencies of using strategies.

## Results

### Data Analysis

The data analysis is carried out using SPSS, version 24. All tests are two-tailed, with an $\alpha$ of .05. In this research it is important to know which strategies the children used to generate their ideas. Therefore, the experimenters asked the children how they came up with their ideas constantly throughout the test. However, one experimenter in the first measurement barely asked this question due to inadequate training. We hence decided to exclude the thirteen children who were tested by this experimenter. Additionally, nine other children, who were not able to finish the test because they did not understand the instructions or were too shy, are also excluded for data analysis, including seven children at the first measurement and two children at the second measurement. In the end, the data of 89 children (40 boys), 3.92-5.00 years old, $M = 4.43$, $SD = 0.28$ in the first measurement, and 51 children (24 boys), 4.33-5.28 years old, $M = 4.78$, $SD = 0.26$, in the second measurement, will be used in this study.

### The Concurrent Influences of Semantic Strategy Use on the Quality of DT

To explore how semantic strategy use influences the quality of DT, four multiple regression analyses were conducted. Since fluency and originality together represent the quality of DT, these analyses includes for each measuring moment an analysis with fluency as dependent variable and an analysis with originality as dependent variable. The percentage

score of semantic strategy use and gender are incorporated as independent variables in all analyses. The data of the whole sample, as presented in table 1, is used.

Table 1

*Descriptive Statistics Alternative Uses Test for the Whole Sample*

| | Boys | | Girls | | Total | |
|---|---|---|---|---|---|---|
| Quality of DT | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Measurement 1[a] | | | | | | |
| Fluency | 18.32 | 10.17 | 19.80 | 8.54 | 19.13 | 9.28 |
| Originality | 7.07 | 7.01 | 7.84 | 6.22 | 7.49 | 6.56 |
| Semantic strategy use | 17.05 | 12.22 | 18.40 | 11.24 | 17.79 | 11.65 |
| Measurement 2[b] | | | | | | |
| Fluency | 18.42 | 6.26 | 21.67 | 7.42 | 20.14 | 7.02 |
| Originality | 11.58 | 7.63 | 15.67 | 11.80 | 13.75 | 10.17 |
| Semantic strategy use | 16.77 | 12.59 | 23.41 | 16.15 | 20.28 | 14.82 |

*Note:* Both measurements have different sample sizes.
[a] N = 89 (40 boys).
[b] N = 51 (24 boys).

Shapiro-Wilk tests revealed that the assumption of no outliers was violated for boys at the first measurement for both the originality, $W(40) = .807$, $p = <.001$, and fluency scores, $W(402) = .885$, $p = <.001$. The corresponding boxplots displayed one obvious outlier, which had great influence on the results because of the small sample size. Hence, we decided to remove this outlier.

The results of the first measurement showed that semantic strategy use and gender explain a significant part of the variances in both fluency, $F(2, 88) = 9.390$, $p = <.001$, $R^2 = .179$, and originality, $F(2, 88) = 10.721$, $p = <.001$, $R^2 = .200$. Regarding the separate predictors, it turns out that the use of semantic strategy is a significant predictor for fluency, $\beta = .417$, $p = <.001$, and originality, $\beta = .444$, $p = <.001$. This means that the more children use the semantic strategy, the higher their DT quality. The results show, on the other hand, no significant relationship between gender and DT quality, neither between gender and fluency, $\beta = -.055$, $p = .574$, nor between gender and originality, $\beta = -.032$, $p = .738$.

Surprisingly, at the second measurement the use of semantic strategy is still a significant predictor of originality, $\beta = .313$, $p = .027$, but no longer for fluency, $\beta = .234$, $p =$

.102. Gender is, on the other hand, still an insignificant predictor for both fluency, β = -.181, *p* = .203, and originality, β = -132, *p* = .344.

Since the results of semantic strategy use are inconsistent over both measurements, it is interesting to know how the strategy use of young children develops over time and whether the strategy use in the first measurement can longitudinally influence the development of their quality of DT. These questions will be answered by the following analyses.

**The Development of Semantic Strategy Use and the Gender Effect**

To investigate the development of semantic strategy use and the effect of gender, a factorial repeated measures ANOVA is used. The percentage score of semantic strategy use is included as dependent variable, and time and gender as independent variables. Since there are only complete data of 35 children (14 boys) for two measurements, 3.94 – 4.88 years old (*M* = 4.42, *SD* = 0.29), only these data will be used in this analysis. The descriptive data are presented in Table 2.

Table 2

*Descriptive Statistics Alternative Uses Test for Children with Repeated Measures*

| | Boys (N = 14) | | Girls (N = 21) | | Total (N = 35) | |
|---|---|---|---|---|---|---|
| Quality of DT | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Measurement 1 | | | | | | |
| Semantic strategy use | 16.90 | 14.52 | 18.43 | 12.01 | 17.80 | 12.91 |
| Fluency | 14.36 | 6.86 | 19.05 | 7.90 | 17.12 | 7.74 |
| Originality | 4.57 | 4.18 | 6.90 | 5.08 | 5.94 | 4.81 |
| Measurement 2 | | | | | | |
| Semantic strategy use | 15.05 | 10.04 | 23.19 | 16.80 | 19.84 | 14.88 |
| Fluency | 20.36 | 6.55 | 22.00 | 7.27 | 21.32 | 6.93 |
| Originality | 13.79 | 8.25 | 15.80 | 11.78 | 14.97 | 10.38 |

Shapiro-Wilk tests revealed that the assumption of normal distribution and no outliers was violated for girls at the second measurement for the scores of the use of semantic strategy, *W*(20) = .770, *p* = <.001. The corresponding boxplots displayed two extreme outliers. However, examining these outliers, they seemed to represent genuine results. Therefore, we decided to keep these outliers as part of the data.

The results of the factorial repeated measures ANOVA showed no significant main effect of time on how often the participants used the semantic strategy, $F(1, 32) = 0.196$, $p = .661$.This indicates that, overall, pre-schoolers do not use semantic strategy significantly more often at the second measurement than they did at the first measurement.

Additionally, the results showed no significant main effect of gender on how often the children used the semantic strategy, $F(1,32) = 1.888$, $p = .179$. Consequently, there seem to be no significant differences between boys and girls regarding how often they use semantic strategy.

The interaction between age and gender reveals more about the development that both boys and girls experience regarding semantic strategy use. The interaction graph, see Figure 1, shows an interesting plot. Although the interaction effect is insignificant, $F(1, 30) = 1.007$, $p = .323$, the plot seems to suggest that the gender of the child had different effects on the development of semantic strategy use, in which the girls use semantic strategy on average 4.8% more at the second measurement, the results of the boys demonstrate an average decline of 1.8%. As a result, the difference between boys and girls regarding how often they use semantic strategy has become bigger at the second measurement.

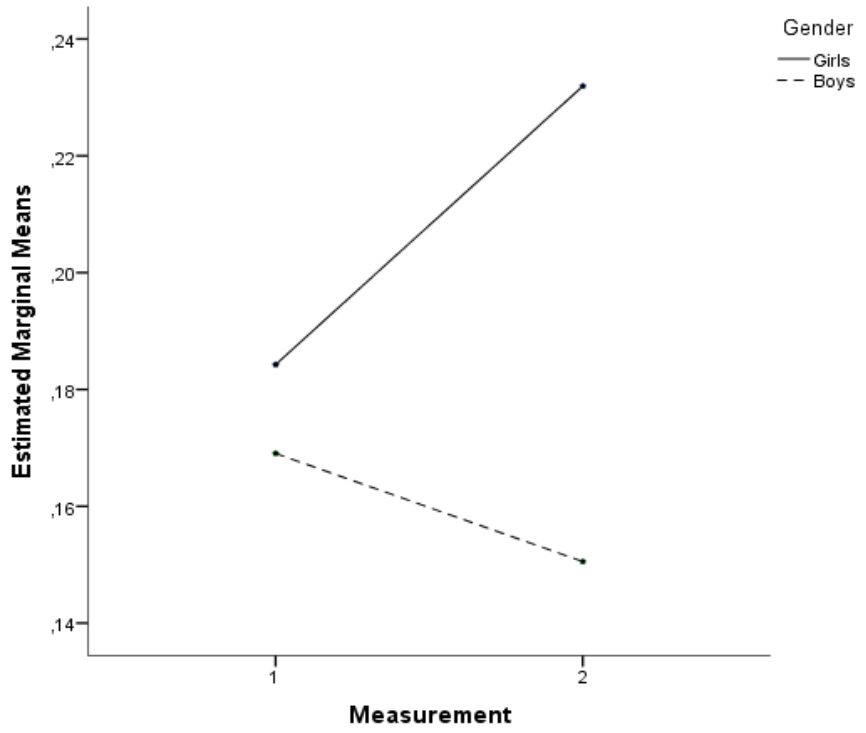


*Figure 1.* Estimated means of semantic strategy use for boys and girls at both measurements.

**The Longitudinal Influences of Semantic Strategy Use and Gender on the Quality of DT**

To investigate how semantic strategy use in the early stage (i.e., at the first measurement) and the gender of pre-schoolers influence their quality of DT over time, a repeated measures MANOVA was conducted. In this analysis, the fluency and originality scores are included as dependent variables, and time, gender and the score of semantic strategy use at the first measurement as independent variables. The data of the same 35 children as shown in Table 2, are used in this analysis.

Shapiro-Wilk tests revealed that univariate normality could only not be assumed for the originality scores for both boys and girls at the first measurement. The corresponding boxplots show positively skewed data with two, although not extreme, outliers. Since these outliers seemed to represent genuine results, we decided to keep these outliers as part of the data.

Findings revealed that the main effect of time is significant, $F(2, 30) = 6.069$, $p = .006$, partial $\eta^2 = 0.399$. This indicates that the overall quality of DT among the children in this sample increases significantly between the first and the second measurement.

Since the longitudinal influence of semantic strategy use and gender on the quality of DT is main interest of this study, we will continue to focus on the interaction effects. The insignificant interaction effect of time and semantic strategy use on the quality of DT, $F(2, 30) = 1.229$, $p = .307$, showed that the semantic strategy use at the first measurement did not influence the development of the quality of DT. The interaction effect of time and gender on the quality of DT is also insignificant, $F(2 ,30) = 0.574$, $p = .569$. This indicates that gender does not influence the development of the quality of DT.

## Discussion

This study was designed to investigate the longitudinal relationship between semantic strategy use and the quality of DT in preschool children. As outlined previously, we formulated two sub questions that contribute to answering the main question.

First of all, we examined the concurrent relationship between semantic strategy use and the quality of DT. The results indicated that semantic strategy use is a significant predictor of DT quality at the first measurement. This is in line with the findings of various adult studies (e.g., Heinonen et al., 2016; Wang et al., 2017). Surprisingly, semantic strategy use is no longer a significant predictor at the second measurement. This could be due to the relatively small sample size of the second measurement (N = 51). A small sample size can limit the magnitude of the correlation, because it is a less reliable representation of the population (Strube, 1991). In addition, the different results could also be due to the fact that

the quality of DT and semantic strategy use seem to develop differently, which results in a changing relationship between both factors over time.

Secondly, we investigated how semantic strategy use develops in young children and whether gender influences this development. Although we expected that children will use advanced strategies more often when they get older (Alexander, et al., 1998; Gilhooly, et al., 2007), the increase in semantic strategy use turned out to be insignificant. One possible explanation for our findings is the relatively narrow age range. A period of five months is very short to measure a child's development. Additionally, the suspected gender differences in semantic strategy use, which we based on recent neuroimaging studies (e.g., Abraham, et al., 2014; Shi, et al., 2017), appeared to be insignificant in this study. Also in this case the small sample size (N = 35) could be a plausible explanation.

Finally, based on the results of the first two analyses, we further examined whether there is a longitudinal influence of the use of semantic strategy on the quality of DT while taking gender into consideration. The findings of this study revealed that the quality of DT increases significantly between the first and the second measurement. This is in line with the research regarding this topic to date. Although researchers did not reach a consensus about the development of DT quality during primary school, most of them agree with an initial increase of DT quality in preschool (e.g., Bijvoet-Van den Berg & Hoicka, 2014, Claxton, et al., 2005). Regarding the gender differences in the quality of DT, insignificant differences were found, despite girls tend to score higher on DT quality. This result is consistent with some previous studies (e.g., Pilar Matud, et al., 2007; Reese, Lee, Cohen, & Puckett, 2001), but inconsistent with a great deal of recent studies that indicate higher DT scores for girls (e.g., Dudek, Strobel, & Runco, 1993; Kuhn & Holling, 2009). These inconsistent results could possibly be explained by different age ranges and school types across the various studies.

With regard to the main question, these findings altogether seem to suggest that the quality of DT and semantic strategy use develop differently. Consequently, semantic strategy use at the first measurement did not influence the development of the quality of DT. It could therefore be concluded that there is no clear longitudinal relationship between DT quality and semantic strategy use.

Yet this explorative study could improve our knowledge about the development of DT in preschoolers. Although this study clearly has some limitations, it can be a good starting point for further research. Additional studies on the current topic with a greater sample size and greater age range are recommended, so to increase the external validity of the results. In the end, these studies could help improving educational practice. Knowing the development

of DT in pre-schoolers, teachers could use this knowledge to teach their children the 21st century skills in a proper way. This will help the children to develop the skills they will need in this rapidly changing society.

## References

Abraham, A. (2016). Gender and creativity: An overview of psychological and neuroscrientific literature. *Brain Imaging and Behavior, 10,* 609-618. doi:10.1007/s11682-015-9410-8

Abraham, A., Thybusch, K., Pieritz, K., & Hermann, C. (2014). Gender differences in creative trhinking: Behavioral and fMRI findings. *Brain Imaging and Behavior, 8,* 39-51. doi:10.1007/s11682-013-9241-4

Alexander, P. A., Graham, S., & Harris, K. R. (1998). *A perspective on strategy research: Progress and prospects, 10,* 129-154. doi:10.1023/A:1022185502996

Baer, J., & Kaufman, J. C. (2008). Gender differences in creativity. *The Journal of Creative Behavior , 42,* 72-105. doi:10.1002/j.2162-6057.2008.tb01289.x

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 42,* 1186-1197. doi:10.3758/s13421-014-0428-8

Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., & Neubauer, A. C. (2014a). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. *NeuroImage., 88,* 125-133. doi:10.1016/j.neuroimage.2013.11.021

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

Bijvoet-Van den Berg, S., & Hoica, E. (2014). Individual differences and age-related changes in divergent thinking in toddlers and preschoolers. *Developmental Psychology, 50,* 1629-1639. doi:10.1037/a0036131

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care, *Assessment and teaching of 21st century skills.* (pp. 17-66). Dordrecht: Springer.

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

Domino, G., & Domino, M. L. (2006). *Psychological Testing. An Introduction.* Cambridge: Cambridge University Press.

Dudek, S. Z., Strobel, M. G., & Runco, M. A. (1993). Cumulative and proximal influences on

the social environment and children's creative potential. *The Journal of Genetic Psychology, 154*, 487-499. doi:10.1080/00221325.1993.9914747

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

Gardner, H. (1982). *Art, mind, and brain: A cognitive approach to creativity*. New York: Basic Books.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *The Britisch Psychological Society*(98), 611-625. doi:10.1111/j.2044-8295.2007.tb00467.x

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

Han, K.-S. (2003). Domain-specificity of creativity in young children: How quantitative and qualitative data support it. *Journal of Creative Behavior, 37*, 117-142. doi:10.1002/j.2162-6057.2003.tb00829.x

Hass, R. W. (2017). Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. *Memory & Cognition, 45*, 233-244. doi:10.3758/s13421-016-0659-y

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, 11*(9), 1-16. doi:10.1371/journal.pone.0162234

Jung, R. E., & Ryman, G. (2013). Imaging Creativity. In K. H. Kim, J. C. Kaufman, J. Baer, & B. Sriraman, *Creatively Gifted Students are not like Other Gifted Students: Research, Theory, and Practice.* (pp. 69-88). Rotterdam: Sense Publishers.

Kershner, J. R., & Ledger, G. (1985). Effect of sex, intelligence, and style of thinking on creativity: A comparison of gifted and average IQ children. *Journal of Personality and Social Psychology*(48), 1033-1040. doi:10.1037/0022-3514.48.4.1033

Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal, 18*, 3-14. doi:10.1207/s15326934crj1801_2

Kim, K. H. (2011). The APA 2009 division 10 debate: Are the Torrance Tests of Creative Thinking still relevant in the 21st century? *Psychology of Aesthetics, Creativity, and the Arts, 5*, 302-308. doi:10.1037/a0021917

Kogan, N. (1974). Creativity and sex differences. *The Journal of Creative Behavior, 8*, 1-14. doi:10.1002/j.2162-6057.1974.tb01103.x

Kuhn, J.-T., & Holling, H. (2009). Exploring the nature of divergent thinking: A multilevel

analysis. *Thinking Skills and Creativity., 4*, 116-123. doi:10.1016/j.tsc.2009.06.004

Larson, L. C., & Miller, T. N. (2011). 21st century skills: Prepare students for the future. *Kappa Delta Pi Record, 47*, 121-123. doi:10.1080/00228958.2011.10516575

Nusbaum, E. C., & Silvia, P. J. (2011). Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence., 39*, 36-45. doi:10.1016/j.intell.2010.11.002

Pilar Matud, M., Rodríguez, C., & Grande, J. (2007). Gender differences in creative thinking. *Personality and Individual Differences, 43*, 1137-1147. doi:10.1016/j.paid.2007.03.006

Razumnikova, O. M., Volf, N. V., & Tarasova, I. V. (2009). Strategy and results: Sex differences in electrographic correlates of verbal and figural creativity. *Human Physiology, 35*, 285-294. doi:10.1134/S0362119709030049

Reese, H. W., Lee, L.-J., & Cohen, S. H. (2001). Effects of intellectual variables, age, and gender on divergent thinking in adulthood. *International Journal of Behavioral Development, 25*, 491-500. doi:10.1080/01650250042000483

Reiter-Palmon, R., & Arreola, N. J. (2015). Does generating multiple ideas lead to increased creativity? A comparison of generating one idea vs. many. *Creativity Research Journal, 27*, 369-374. doi:10.1080/10400419.2015.1087274

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal, 24*, 66-75. doi:10.1080/10400419.2012.652929

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*, 92-96. doi: 10.1080/10400419.2012.650092

Russo, C. F. (2004). A comparative study of creativity and cognitive problem-solving strategies of high-IQ and average students. *Gifted Child Quarterly, 48*, 179-190. doi:10.1177/001698620404800303

Shi, B., Xu, L., Chen, Q., & Qiu, J. (2017). Sex differences in the association between gray matter volume and verbal creativity. *NeuroReport, 28*, 666-670. doi:10.1097/WNR.0000000000000820

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*, 104-109. doi:10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Jenkins, E. (2014). *How children discover new strategies.* New York: Psychology Press.

Strube, M. J. (1991). Demonstrating the influence of sample size and reliability on study

outcome. *Computers in Teaching, 18,* 113-115. doi:10.1207/s15328023top1802_15

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. doi:10.1016/j.neuropsychologia.2017.03.001

Zeng, L., Proctor, R. W., & Salvendy, G. (2011). Can traditional divergent thinking tests be trusted in measuring and predicting real world creativity? *Creativity Research Journal, 23,* 24-37. doi:10.1080/10400419.2011.545713

## Appendix I

Overview of the basic categories and the grouped categories.

| Grouped category | Category | Definition | Type of basic category |
|---|---|---|---|
| Memory based production | Dominant use | States the main use of a stimulus without explanation | Use-related category |
| | Unmediated use | State possible use without explanation | Use-related category |
| | Episodic memory use | States possible use with reference to a specific memory | Use-related category |
| | Experience-derived use | Generate a use that is derived from prior knowledge | Use-related category |
| Abstraction/ substantiation based production | General use | States very wide category of use | Use-related category |
| | Example use | After giving a higher abstract level use, subsequently gives more specific examples (with lower abstract levels) | Use-related category |
| | Broad use first | Considers target stimulus against an initially broad use category | Use-related category |
| | Later general use | After generating one or more specific uses, propose a more abstract use that can include the previous use(s) as its example(s) | Use-related category |
| Wholeness breaking based production | Disassembly uses | States a way of decomposing the target stimulus and using the resulting components | Use-related category |
| | Assembly use | Assemble a stimulus or parts of the stimulus with (an)other new object(s) to generate a use | Use-related category |
| Property based production | Property use | Explicitly indicates a property/properties that enables the stated use for the target item | Use-related category |
| | Analogy use | Bases on the similarity of different things (based on the features but not the action) to generate a use | Use-related category |
| | Property | States a property of the stimuli | Use-unrelated categories |
| | Environment-based use | Children are consciously or unconsciously influenced by | Use-related categories |

| | | | |
|---|---|---|---|
| Environment based production | | the test environment and generate a use | |
| | Environment | Children describe the test environment without generating a use | Use-unrelated categories |
| Imagination based production | Imagination use | Uses imagination to generate a use | Use-related categories |
| Self-cueing | Context | Mentions the context in which the target stimulus is often found | Use-unrelated categories |
| | Imagery | Indicates forming a mental image | Use-unrelated categories |
| | Item naming | Repeats the name of a stimulus | Use-unrelated categories |
| | Use query | Poses use problem | Use-unrelated categories |
| Impasse | Repeat use | Repeats an already stated use for a target stimulus | Exceptional categories |
| | Impasse | Indicated participant cannot report any further uses at this point | Exceptional categories |
| Self-transformation | Self-transformation | The child transformed learned rules/methods/strategies to new situations | Use-unrelated categories |
| Self-reflection | Self-reflection | Questions self about the validity of generated use | Use-unrelated categories |
| Meta evaluation of the task | Meta evaluation of the task | Children evaluate the task | Use-unrelated categories |

**Definitions of the grouped categories**

| **Grouped category** | **Definition** |
|---|---|
| Memory based production | Generating uses from memory, either with or without reference to prior knowledge. |
| Abstraction/ substantiation based production | Generating uses via using the conceptual connection, which includes two processes: abstraction and substantiation. |
| Wholeness breaking based production | Generating uses based on breaking the mind-set of viewing a stimulus as a whole thing; a stimulus can either be broken down into pieces or be a part of another thing. |

| | |
|---|---|
| Property based production | Generating uses based on the properties or features of a stimulus. |
| Environment based production | Generating uses via using information in the surroundings. |
| Imagination based production | Generating uses via imagining. |
| Self-cueing | Processes to recall information that relate to the target stimulus. |
| Impasse | Repeating generated uses, or fall in obvious impasse indicates the process of impasse of generating more uses. |
| Self-transformation | Transforming learned rules/methods/strategies to new situations by the child him- or herself. |
| Self-reflection | Evaluating the validity of generated uses by him- or herself. |
| Meta evaluation of the task | Evaluating the task. |