

Analysis of young children's abilities to cluster and switch during a verbal fluency task

Citation for published version (APA):

Resch, C., Martens, R., & Hurks, P. (2014). Analysis of young children's abilities to cluster and switch during a verbal fluency task. *Neuropsychology, Development and Cognition. Section D: The Clinical Neuropsychologist*, 28(8), 1295-1310. <https://doi.org/10.1080/13854046.2014.978382>

Document status and date:

Published: 01/01/2014

DOI:

[10.1080/13854046.2014.978382](https://doi.org/10.1080/13854046.2014.978382)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

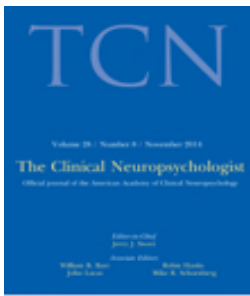
www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.



Analysis of young children's abilities to cluster and switch during a verbal fluency task

Christine Resch, Rosa Martens & Petra Hurks

To cite this article: Christine Resch, Rosa Martens & Petra Hurks (2014) Analysis of young children's abilities to cluster and switch during a verbal fluency task, The Clinical Neuropsychologist, 28:8, 1295-1310, DOI: [10.1080/13854046.2014.978382](https://doi.org/10.1080/13854046.2014.978382)

To link to this article: <https://doi.org/10.1080/13854046.2014.978382>



Published online: 13 Nov 2014.



Submit your article to this journal [↗](#)



Article views: 254



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Analysis of young children's abilities to cluster and switch during a verbal fluency task

Christine Resch, Rosa Martens, and Petra Hurks

Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

To investigate developmental changes that take place in verbal fluency (VF) performance during early childhood, a VF task was administered to 225 healthy, Dutch-speaking children aged between 4.14 and 6.89 years. Three categories of VF outcome measures were included: i.e., word productivity, mean cluster size, and number of switches. Age influenced performance on all VF outcome measures linearly; i.e., older children produced more words, made longer clusters, and switched more. Higher levels of intelligence were associated with increased VF word productivity, but not with measures of switching and clustering. When leaving intelligence out of these analyses, we additionally found an interaction between level of parental education (LPE) and sex on total word productivity, i.e., girls with parents who had lower LPE produced fewer words than the other children. Furthermore, a similar interaction of LPE and sex was found for the number of switches: i.e., girls who had parents with lower LPE made fewer switches than the other children. Findings suggest that even in 4 to 6-year-old children important changes take place over time in VF and in processes underlying successful performance. Attention should be paid to age-extrinsic factors, such as LPE and sex, since these have been found to influence VF performance in young children.

Keywords: Verbal fluency; Development; Clustering; Switching.

INTRODUCTION

Verbal fluency (VF) tasks are widely used in clinical practice as well as in research settings (Lezak, Howieson, Bigler, & Tranel, 2012; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006). In VF tasks, the goal is to name as many words as possible from a certain category (e.g., animals, food, words beginning with the letter "S") within one minute (Troyer, Moscovitch, & Winocur, 1997). In general, VF tasks are believed to be valid tools for measuring a variety of cognitive abilities, such as lexicosemantic knowledge, lexical access (Sergeant, Geurts, & Oosterlaan, 2002), and executive functioning, including cognitive flexibility, monitoring, and strategy use (Hurks et al., 2010; Troyer et al., 1997; Van der Elst, Hurks, Wassenberg, Meijs, & Jolles, 2011). VF task performance has been linked to frontal lobe functioning (Baldo, Schwartz, Wilkins, & Dronkers, 2006; Romine & Reynolds, 2005; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998) as well as to temporal lobe functioning (Troyer et al., 1998).

Address correspondence to: Department of Neuropsychology and Psychopharmacology, Maastricht University, P. O. Box 616 NL – 6200 MD, Maastricht, The Netherlands. E-mail: pm.hurks@maastrichtuniversity.nl
(Received 6 February 2014; accepted 14 October 2014)

In the last years, VF tasks have been increasingly used as a tool in pediatric neuropsychological research (Hurks, 2012). For instance, researchers have found that VF performance (i.e., total words generated over 60 seconds) improves with age, as studied in samples of children aged 6 to 16 (Hurks et al., 2010; Kavé, Kigel, & Kochva, 2008; Sauzéon, Lestage, Raboutet, N’Kaoua, & Claverie, 2004; Tallberg, Carlsson, & Lieberman, 2011; Hurks, 2012). Surprisingly, however, children younger than 6 years of age were rarely included in these studies (Hurks et al., 2010; Kavé et al., 2008; Koren, Kofman, & Berger 2005; Sauzéon et al., 2004; Tallberg et al., 2011). To be able to obtain a more complete overview of the developmental changes that take place in VF performance during childhood, the present study aimed to investigate VF performance of children as young as 4–6 years of age in more depth.

The few studies that have investigated VF performance in children younger than 6 years found that the total number of words generated over 60 seconds increased in children from age 3 onwards (Ardila, Rosselli, Matute, & Guajardo, 2005; Klenberg, Korkman, & Lahti-Nuutila, 2001; Romine & Reynolds, 2005). Furthermore, previous studies showed that a large developmental spurt in executive functioning (i.e., the cognitive functions underlying VF performance) can be observed between age 5 and at least age 8 (Korkman, Lahti-Nuutila, Laasonen, Kemp, & Holdnack, 2013; Romine & Reynolds, 2005).

Unfortunately, the above-mentioned studies on early childhood VF performance (Ardila et al., 2005; Klenberg et al., 2001; Romine & Reynolds, 2005) only took into account the classical quantitative method of scoring VF performance, namely counting the number of correct words generated over 60 seconds. As described previously, VF tasks tap into a variety of cognitive functions, among which are lexico-semantic knowledge, cognitive flexibility, and self-monitoring (Hurks et al., 2010; Troyer et al., 1997; Sergeant et al., 2002; Van der Elst et al., 2011). The classical scoring method does not allow a differentiation between these underlying cognitive processes. Past research has shown that by looking at the systematic organization of information (i.e., clustering and switching), valuable insights can be obtained with regard to underlying processes involved in VF (Hurks et al., 2004, 2006; Robert et al., 1998; Troyer, 2000; Troyer et al., 1997).

Clustering is a reflection of the ability to recall three or more associated words together (Robert et al., 1998; Troyer, 2000; Troyer et al., 1997). It can be quantified by calculating the mean size of the clusters generated within 60 seconds (Hurks et al., 2010; Kavé et al., 2008; Sauzéon et al., 2004; Tallberg et al., 2011). In an animal VF task with the objective to name as many animals as possible, a cluster can, for example, consist of “insects” or “birds”. The ability to cluster words involves accessing and using words from memory and is generally seen as a measure of lexico-semantic knowledge (Troyer et al., 1997, 1998).

Switching, on the other hand, can be defined as the ability to initiate clustering and to switch to new categories (Hurks, 2012; Hurks et al., 2010; Troyer et al., 1997). It involves search processes that are assumed to depend on executive functions, such as response initiation, monitoring, set shifting, and cognitive flexibility (Hurks et al., 2010; Kavé et al., 2008; Troyer et al., 1997). To assess switching, the total number of switches made between clusters, between clusters and single words, and between single words has to be counted (Hurks et al., 2010; Kavé et al., 2008; Troyer et al., 1997). However, since the calculation of the number of switches also includes switches to and

among single words, it has been suggested that this score is actually an indication of an inability to cluster, rather than a reflection of an executive shifting process (Abwender, Swan, Bowerman, & Connolly, 2001; Koren et al., 2005). To overcome this problem, an alternative measure of switching which does not include single words can be used, i.e., the number of clusters (Hurks et al., 2010; Kavé et al., 2008; Koren et al., 2005). By calculating the number of clusters, an indication can be given of the ability to initiate the use of an associative strategy (Hurks et al., 2010; Koren et al., 2005). The number of clusters can therefore be seen as a useful addition for evaluating strategy use in VF performance, and as an essential component of alternative qualitative analysis of VF performance next to the number of switches and the mean cluster size.

To validate the use of these alternative VF scoring methods, researchers have compared results of a VF task to results of, for example, a design fluency (DF) task (Abwender et al., 2001; Hurks et al., 2010). The DF task is a non-verbal fluency test with the objective to draw as many abstract designs as possible within 60 seconds (Van der Elst et al., 2011). Since performance on a DF task is also thought to depend strongly on strategy use (Lezak et al., 2012), this task has been used to assess the construct validity of the VF outcome measures of clustering and switching. Research comparing quantitative and qualitative performance components of VF tasks to performance on a DF task found that correlations between these two tasks were high (Abwender et al., 2001; Hurks et al., 2010). Furthermore the developmental trajectories described above for VF clustering and switching components are highly comparable to those seen in the DF task (Hurks et al., 2010). Therefore, it can be concluded that the measures of clustering and switching that are used in the alternative scoring method of the VF task are valid measures of specific cognitive functions, at least in older children.

The measures of clustering and switching have been studied as an alternative scoring method of VF tasks in pediatric neuropsychological research in older children (Hurks, 2012; Hurks et al., 2010; Kavé et al., 2008; Sauzéon et al., 2004; Tallberg et al., 2011). However, to the authors' knowledge, no studies have used this alternative method to evaluate performance of children below the age of 5. The second aim of the present study was therefore to include the alternative scoring method of VF tasks in the assessment of young children's performance. It was expected that these alternative outcome measures (i.e., the mean cluster size, the number of clusters, and the number of switches) would correlate at least to some degree with VF word productivity over 60 seconds, since the cognitive processes that these alternative outcome measures rely on are suggested to be important for optimal overall VF performance (Hurks et al., 2010; Kavé et al., 2008; Koren et al., 2005; Sauzéon et al., 2004; Tallberg et al., 2011).

Studies including older children have led to important insights and results. For one, when studying children between the ages of 6 to 15, Hurks and colleagues (2010) found an increase in the number of clusters until at least 12 to 13 years of age, while the number of switches increased until at least 14 years of age. Similar increases in the number of clusters and switches with age were also reported by several other studies (e.g., Kavé et al., 2008; Koren et al., 2005; Tallberg et al., 2011). Mean cluster size has been reported to increase with age until at least 11 years (Hurks et al., 2010; Kavé et al., 2008; Sauzéon et al., 2004). Based on these findings, one may conclude that cluster size, a measure of lexico-semantic organization, is established earlier in development than switching, a measure of executive functioning. In accordance with studies on VF performance in older children, we hypothesized that, as early as 4 to 6 years of age, an

age-related increase would be seen in the total number of correctly generated words, the mean cluster size, the number of clusters and the number of switches. Since the greatest period of development in executive functions is reported to be between the ages of 5 and 8 (Korkman et al., 2013; Romine & Reynolds, 2005), it was expected that changes in these functions, as measured by the various qualitative components of the VF task, could already be detected in children between age 4 and 6.

Next to analyzing quantitative and qualitative scores of the VF task, the current study aimed to contribute to the growing body of research that has reported significant effects of several child-related factors, such as age, sex, and intelligence, on changes in VF performance (Ardila et al., 2005; Hurks et al., 2010; Klenberg et al., 2001; Prigatano, Gray, & Lomay, 2008; Van der Elst et al., 2011). Past studies on VF performance in primary school children have often not included such age-extrinsic factors, even though they are thought to explain subtle differences in cognitive development (Hurks et al., 2010; Klenberg et al., 2001). The present study included such factors, i.e., age, level of parental education (LPE), sex, and intelligence, to achieve an even broader view on VF performance at a young age.

In previous studies, a higher LPE was found to have a positive influence on VF performance – at least on the total number of words produced over 60 seconds (Ardila et al., 2005; Hurks et al., 2006; Klenberg et al., 2001). When using the above-mentioned alternative scoring methods, Hurks and colleagues (2010) found a positive effect of higher LPE on mean cluster size, while there was no influence on the number of clusters or the number of switches in children aged 6 to 16. These results seem to indicate that LPE is positively associated with lexico-semantic knowledge and lexical access, as measured by the mean cluster size, while the influence of LPE on executive functioning, as assessed by the number of clusters and the number of switches, is less. In accordance with previous studies (Ardila et al., 2005; Hurks et al., 2010; Klenberg et al., 2001), it was hypothesized that a higher LPE would be associated with an larger mean cluster size and a higher total number of correct words generated. Additionally, it was expected that a higher LPE would have no effect on the measures of switching (in line with Hurks et al., 2010).

Next, studies on the association between sex and VF performance are still inconclusive (Hurks et al., 2010; Koren et al., 2005; Van der Elst et al., 2006; Weiss et al., 2006). Although many studies did not find an effect of sex on VF performance (Brocki & Bohlin, 2004; Hurks et al., 2006; Lanting, Haugrud, & Crossely, 2009; Riva, Nichello, & Devoti, 2000; Van der Elst et al., 2011), other studies have reported significant results in this field. Some studies found that boys outperform girls on VF tasks (Ardila et al., 2005; Prigatano et al., 2008), while others reported that girls perform better than boys (Klenberg et al., 2001). Given the inconsistent past findings regarding the effect of sex on VF performance, this variable was included in the present study to examine its influence on the various components of VF tasks in young children.

A third age-extrinsic factor was taken into account in the present study, i.e., intelligence quotient (IQ). Previous research in children has found that children's intelligence level is positively associated with both the total number of words generated on a VF task (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001) and tests of executive functions (Mahone et al., 2002). No studies have been conducted regarding the relationship between intelligence and VF clustering and switching. However, based on

studies including tests of executive function other than the VF task, we hypothesized that intelligence has a positive influence on VF outcome measures of total number of correct words, number of clusters and number of switches.

In sum, the primary goal of the present study was to investigate VF performance in a sample of healthy 4 to 6-year-old children. Next to the quantitative outcome measure of the total number of correctly generated words, three alternative qualitative VF scores were included, i.e., mean cluster size, number of clusters, and number of switches. To our knowledge, this was the first study that has applied this alternative scoring method of VF tasks in a sample of children as young as age 4–6 years. Age-extrinsic factors (i.e., sex, LPE, and intelligence) were also considered when examining the development of VF performance in young children, since the influence of these factors might contribute to the observed inconsistencies in results of previous studies.

METHOD

Procedure and participants

All children enrolled in grade 1 or 2 of 24 Dutch primary schools were invited to participate in the present study. In The Netherlands, all children aged 4 years and older have to attend primary school, which is a compulsory form of education. In grades 1 and 2 of this primary school, the children do not yet receive formal instructions in reading, writing and/or mathematics. Caregivers of the children were given an information package, distributed via the schools, and were asked to give written consent for the participation of their child in the study. In addition, they were requested to fill in a questionnaire about the development and medical history of their child, and about their own educational background.

In total, 24% of all caregivers responded to the invitation for participation in the study. After consent of the caregiver was obtained, the children were screened based on exclusion criteria that might have an influence on test performance, i.e., not speaking Dutch fluently, the presence of neurological disorders (e.g., absence epilepsy) or the use of medication (e.g., antihistamines). In The Netherlands, healthcare professionals are hesitant to use DSM labels to classify the behavioral patterns of young children (i.e., children younger than 7 years). Therefore, we expected that the percentage of children classified as having a DSM diagnosis would be very low in our sample. Indeed, during a one year follow-up of the children included in our sample, Martens (2012) found that only five (i.e., 2%) had received a DSM-IV diagnosis (i.e., dyslexia or attention deficit hyperactivity disorder). During this follow-up, these five children were still attending regular primary schools and were in the appropriate grade for their age. Based on these follow-up data, we decided not to include these DSM-IV diagnoses as exclusion criteria in the present study.

The final sample consisted of 225 children (110 boys) aged 4.14 to 6.89 years ($M = 5.16$, $SD = 0.61$). As mentioned above, caregivers completed a questionnaire in which they, among others, had to indicate their highest completed level of parental education (LPE), ranging from primary school (1) to university degree (8) (De Bie, 1987; Kalff et al., 2001). This scaling method is comparable to the International Standard Classification of Education (UNESCO, 1997). Next, LPE scores obtained in the present study were recoded into low (1 and 2), mid (3, 4, and 5), and high (6, 7, and 8) levels

of education (in line with Martens, 2012). If the LPE differed between the mother and the father, the highest level of education was chosen. Of all children included in our sample, 6% could be classified as having caregivers with a low LPE, 37% as those with caregivers having a mid-LPE, and 57% as those having caregivers with a high LPE. The average LPE in the current sample was slightly higher than the (estimated) distribution of educational levels in the Dutch adult population. Indeed, data published in 2012 by the *Centraal Bureau voor de Statistiek* [Bureau of Statistics] in The Netherlands indicates that 35% of Dutch adults can be classified as having a low educational level, 40% as mid-level, and 32% as high level. In our sample the LPE was identical for the mother and the father in 130 cases (58%). For 82 children (36%), the LPE was different for the mother and the father and in 77 of these cases (34% of the total sample), the LPE differed no more than one category (i.e., LPE low vs. mid, or LPE mid vs. high). In 40 cases (18%), the mother had a higher LPE than the father and in 42 cases (19%) the father had a higher LPE than the mother. For 9 children (4%), the LPE used was obtained from the mother due to an unknown LPE of the father, and for 2 children (1%), the LPE of the father was used since the LPE of the mother was unknown. Lastly, for 2 children (1%), information about LPE was lacking for either caregiver.

The Ethics Committee of the Faculty of Psychology of Maastricht University approved the research protocol. Well-trained research assistants administered the semantic VF task and the Raven's CPM, as a measure of intelligence, in the same order for each child in a stimulus-free room at the participating schools.

Measures

Verbal fluency (VF) task. Participants were required to generate as many animals as possible. This variant is frequently used (Hurks, 2012; Tallberg et al., 2011; Van der Elst et al., 2011), which is helpful when comparing the results to research done in other age groups or in clinical populations (e.g., children with attention deficit hyperactivity disorder, see Hurks et al., 2004). In the present study, the main outcome measure was the total number of correctly generated words over 60 seconds. In line with Troyer (2000), this outcome measure did not include the number of perseverations (e.g., cat and cats) and incorrect words (i.e., non-animal words such as "car", or names of an animal, such as "Sally"). If perseverations and incorrect responses were to be included, the total number of words (i.e., six) of a child who repeated many words (e.g., cat, cats, dog, dogs, cow, cows) would seem higher than the total number of words (i.e., three) of a child who named the same words without repetition (e.g., cat, dog, cow), even though the number of unique words is the same for both children. Therefore, repetitions and incorrect words were not counted when estimating the total number of correctly generated words.

Apart from the total number of correctly generated words, the mean cluster size, the number of switches, and the number of clusters was calculated. In the present study, a cluster was defined as a group of successively generated words consisting of at least three categorically related words, in line with Troyer et al. (1997) and Robert et al. (1998). Various criteria can be used to define a category. In line with Troyer and colleagues (1997), categories in the current research were based on zoological families (e.g., birds, insects, primates), living environment (e.g., Africa, farm) or human use (e.g., pets).

Mean cluster size. Mean cluster size was computed by counting all the words that are related by clusters and dividing them by the number of clusters (Troyer et al., 1997). For example, in the following list – *walrus, fur seal, sea lion, monkey, hen, rooster, goose, whale, fly, cockroach, beetle, snake, pigeon, seagull, owl, canary* – the number of words related by cluster is 13 (list adapted from Hurks et al., 2010), and the number of clusters is four. Therefore, the mean cluster size in this example is 13 divided by 4. Perseverations, as defined earlier, were included when calculating the number of words related by cluster, and the number of clusters (in line with Troyer, 2000). In general, children do not name the same word twice in a row. However, an animal can belong to different categories, for example, a *rooster* can belong in the category *farm animals* while it can also be a part of the category *birds*. Therefore, it is possible that a child first names a number of farm animals, including *rooster* (e.g., pig, *rooster*, cow, horse), and later on names the same animal while listing animals from a different category (e.g., canary, eagle, *rooster*, pigeon). By not including these perseverations, mean cluster size or the number of clusters (since only two words belonging to the same category would not meet the criteria of a minimum of three words for a cluster) would decrease while the words were actually named as part of a cluster. To obtain a reliable measure of the qualitative VF outcomes, it is therefore more important to recognize this clustering ability than to exclude the repetition of a word.

The number of clusters. This was calculated by counting the number of clusters generated within one minute. For example, in the example list from Hurks et al. (2010), four clusters can be identified: sea animals, birds, insects, and birds. Perseverations, as defined earlier, were included in the clusters (in line with Troyer, 2000).

The number of switches. This was defined as the number of switches made between two clusters, between a cluster and a single word, or between two single words. For example, in the example listed above, six switches were made. When calculating the number of switches, perseverations and incorrect words were included (in line with Troyer, 2000) since they provide information about how often a child names a single word or begins naming words from a new category.

Raven's Coloured Progressive Matrices (Raven's CPM; Raven, 1958). The CPM is a measure of intelligence consisting of 36 items. The child is shown a pattern with a missing piece and from six possible pieces has to select the piece that completes the pattern. Items gradually increase in difficulty. One point is awarded per correct answer, with a maximum total raw score of 36. The total raw score obtained on this test was used in the analyses. The measure has been shown to have a satisfactory reliability (Lezak et al., 2012).

Statistical analysis

Firstly, bivariate correlation analyses were performed to investigate the degree to which the independent (i.e., age, sex, LPE, and intelligence) and dependent variables (i.e., total number of correct words, mean cluster size, number of clusters and number of switches) were associated. This not only allowed us to assess the strength of the

relationship between independent and dependent variables, for example between age and VF clustering, but also the association among the dependent variables themselves, for example, the VF number of clusters and the number of switches. These potential associations among the variables discussed above were tested using Pearson's correlations. The only exceptions here are those potential associations involving LPE. Due to the categorical nature of LPE (with three categories: low, mid and high), associations involving this variable were tested using Spearman's non-parametric correlation statistics.

Secondly, to study the added effect of one independent variable on VF performance over and above the effects of the other independent variables, multiple linear regression analyses were conducted (in line with Van der Elst et al., 2011; Hurks, 2012). By using this type of analysis, the combined influence of various demographic predictors on the outcome variables can be investigated. First, the predictor variables age, sex (coding: girls = 0, boys = 1), LPE, and score on the Raven's CPM were entered simultaneously in the initial analysis. LPE was dummy coded with two dummies (LPE low and LPE high) and LPE mid as a reference category. Furthermore, all two-way interactions between age, sex, and LPE dummies were entered as predictors in the regression models. To avoid multicollinearity, age was centered before computing the interactions (Marquardt, 1980). For each model, the data were tested for multicollinearity by calculating the Variance Inflation Factors (VIF), which should not exceed 5, and the Condition Indices (CIs), which should not exceed 15 (Kutner, Nachtsheim, Neter, & Li, 2005). These assumptions were met in all regression models. However, since higher VIF values and CIs might bias the regression model (Bowerman & O'Connell, 1990), it was aimed to keep these values as low as possible. Therefore, to achieve small VIF values and CIs, the number of correlated predictors in the model was decreased. Thus, instead of keeping all the entered predictors in the model, the full regression models were reduced in a step-down hierarchical procedure. This step-down procedure has been described as the only appropriate method for theory testing (Studenmund & Cassidy, 1987), since it is not influenced by random variation in the data, thereby making it possible to find replicable results with retesting. More recently, however, Millis (2003) has criticized this method since it increases the chance of coincidentally finding a significant relationship between the dependent variable and independent variables (i.e., Type I error). A good method to avoid this increase in Type I error rate is to perform a Bonferroni-correction for each of the models, taking into account the number of predictors (Mundfrom, Perret, Schaffer, Piccone, & Roozeboom, 2006). The significance-level was therefore determined for each model by dividing the nominal α -level of .05 by the number of predictors in the model (Mundfrom et al., 2006). For example, for a model with 5 predictors, the α -level was $.05/5 = .010$. Since computer-controlled stepwise procedures often produce results that are difficult to replicate (Menard, 2002), the predictor that was found most non-significant was manually excluded from the model. This procedure was repeated until only significant predictors were left. Note, however, that a predictor was not deleted from the model as long as it was part of a higher-order predictor in the model to avoid arbitrariness of the p -value (Aiken & West, 1991). For example, age was never removed if an interaction involving age was still in the model (in line with Van der Elst et al., 2006). Since the LPE dummies (LPE low and LPE high) both represent the effect of the categorical predictor LPE, they were always kept in or deleted from the model together. Furthermore, the

interactions of the LPE dummies were also either both included or excluded from the model (in line with Van der Elst et al., 2006).

All analyses that included the Raven's CPM scores were based on the raw scores of 223 children, since the administration of this task could not be completed for two children in the sample, and for a further two children, information about LPE was missing. Therefore, all analyses including LPE were also based on the data of 223 children. Analyses including both the Raven's CPM and LPE were based on the data of 221 children.

RESULTS

Correlations among VF outcome measures

In terms of correlations among the VF measures, the total number of correct words correlated positively with the mean cluster size, the number of clusters, and the number of switches, showing that optimal overall VF performance depends on a variety of cognitive processes. The number of clusters correlated positively with the mean cluster size. This suggests that children who are able to initiate clustering successively generate more associated words (i.e., produce larger mean cluster size). A significant negative correlation was found between the mean cluster size and the number of switches, indicating that the longer the clusters, the less switches take place. The number of clusters and the number of switches were not correlated, implying that even though they are both thought to be measures of executive functions, they are dependent on unrelated cognitive functions. A complete overview of the correlations can be found in Table 1.

Influence of sex, LPE, and intelligence on VF performance

Total word productivity. An overview of the results of the final regression analyses is presented in Table 2. In terms of total VF scores, Raven's CPM, $t(222) = 3.62$, and age, $t(222) = 4.12$, were found to be significant predictors of the number of correct words generated within 60 seconds. In other words, higher scores in terms of intelligence and a higher age were both associated with higher VF word productivity.

Table 1. Correlations between age, sex, level of parental education, Raven's CPM score and verbal fluency measures

Variables	1	2	3	4	5	6	7	8
1. Age	–							
2. Sex	.04	–						
3. LPE	-.14*	-.07	–					
4. Raven's CPM	.49***	.06	.14*	–				
5. Total correct words	.41***	-.11	.00	.39***	–			
6. Mean cluster size	.21**	.02	-.01	.16*	.42***	–		
7. Number of clusters	.29***	.12	-.01	.20**	.65***	.6***	–	
8. Number of switches	.25***	.03	-.04	.24***	.65***	-.21**	.07	–

LPE = Level of parental education. Raven's CPM (Raven's Coloured Progressive Matrices; Raven, 1958) = a measure of intelligence.

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

Table 2. Final multiple linear regression models of the verbal fluency measures

	Predictor	B	SE B	β	R^2
Total correct words	<i>Raven's CPM</i>	0.17	0.05	0.25***	0.21*** ^a
	<i>Age</i>	1.48	0.36	0.28***	
Total correct words (excluding Raven's CPM from the analyses)	<i>Age</i>	2.17	0.32	0.42***	0.22*** ^b
	<i>Sex</i>	-0.55	0.63	-0.09	
	<i>Low LPE</i>	-3.07	1.17	-0.24**	
	<i>High LPE</i>	0.43	0.58	0.07	
	<i>Sex by Low LPE</i>	4.64	1.64	0.26**	
	<i>Sex by High LPE</i>	-0.55	0.81	-0.08	
Mean cluster size	<i>Age</i>	0.73	0.23	0.21**	0.05** ^c
Number of clusters	<i>Age</i>	0.38	0.08	0.29***	0.09*** ^d
Number of switches	<i>Age</i>	1.09	0.29	0.25***	0.10** ^c
	<i>Sex</i>	-0.30	0.57	-0.06	
	<i>Low LPE</i>	-1.94	1.05	-0.18	
	<i>High LPE</i>	0.09	0.52	0.02	
	<i>Sex by Low LPE</i>	4.06	1.48	0.27**	
	<i>Sex by High LPE</i>	-0.12	0.73	-0.02	

LPE = Level of parental education. Raven's CPM (Raven's Coloured Progressive Matrices; Raven, 1958) = a measure of intelligence.

^a $F(2, 222) = 29.62, MSE = 7.90$, ^b $F(6, 216) = 10.19, MSE = 8.07$, ^c $F(1, 223) = 10.49, MSE = 4.16$, ^d $F(1, 223) = 20.99, MSE = 0.58$, ^e $F(6, 216) = 3.84, MSE = 6.55$.

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

When the intelligence measure (i.e., Raven's CPM) was left out of the regression models, results changed. More specifically, next to age, $t(219) = 6.83$, the dummy variable for low LPE, $t(219) = -2.63$ and the interaction between the low LPE dummy and sex, $t(219) = 2.82$, became significant. Together, these results seem to indicate that children are negatively affected by low LPE, in that they produce less correct words when their parents are lower educated. However, this effect seems to apply only to girls.

Mean cluster size, number of clusters and number of switches. While examining the qualitative VF scores (i.e., clustering and switching), mean cluster size could only be predicted linearly by age, $t(223) = 3.24$. Older children produced more words belonging to a cluster. Age was also found to be the only predictive variable for the number of clusters, $t(222) = 4.58$, meaning that the number of clusters increases as the children grow older.

Finally, with regard to the number of switches, age, $t(219) = 3.80$, was found to be of significant predictive value. Furthermore, a sex by low LPE interaction, $t(219) = 2.74$, was associated with the number of switches, in that girls with lower educated parents make fewer switches.

DISCUSSION

The aim of the present study was to investigate VF performance in 4 to 6-year-old children using both a traditional, quantitative measure (i.e., total number of correct words) and an alternative qualitative scoring method (i.e., mean cluster size, number of clusters and number of switches). Qualitative VF outcome measures are thought to be

of great value for investigating the cognitive processes that underlie VF task performance, such as lexico-semantic knowledge, cognitive flexibility and strategy use. Since early childhood (i.e., younger than age 7) is a period where a developmental spurt occurs in many of these cognitive functions (Korkman et al., 2013; Romine & Reynolds, 2005), an increase in performance on both quantitative and qualitative VF measures with age was expected. Additionally, the effects of sex, LPE, and intelligence were examined to discover the influence of these factors on VF performance in young children.

First, we found that the mean cluster size, the number of clusters and the number of switches were positively related to the total number of correct words produced on the VF task (in line with studies that include older children, e.g., Kavé et al., 2008; Koren et al., 2005). The correlations between the qualitative VF outcome measures themselves were found to be (only) moderately high. Taken together, these results indicate that at age 4–6 years, fluency performance already depends on various functions, such as lexico-semantic knowledge, cognitive flexibility, and use of word association strategies. This is an important finding, since the present study shows that next to lexico-semantic knowledge, strategy use already has a beneficial influence on VF task performance from the age of 4 onwards. Thus, even in very young children, VF tests can already provide valuable information about development of various cognitive functions. Furthermore, compared to younger, 4-year-old children, older, 6-year-old children generated more words during a VF task. Additionally, they clustered more words, named more words in one cluster (i.e., the mean cluster size increased), and switched more than 4-year-old children. Increases in these VF outcome measures between the ages of 4 and 6 years showed that the functions underlying VF performance are already developing at an early age.

Next, our findings revealed that two of the VF outcome measures, i.e., total number of correct words and number of switches, were also influenced by age-extrinsic factors. Analyses showed that VF total word productivity (i.e., a quantitative VF measure) was positively associated with intelligence. This is in line with previous studies indicating that a higher intelligence level has a positive influence on quantitative VF performance (Anderson et al., 2001). In the present study, Raven's CPM was chosen as the measure for intelligence. This instrument has been generally acknowledged as a valid and reliable measure of fluid intelligence (Cherkes-Julkowski, Stolzenberg, & Segal, 1990). Fluid intelligence can be defined as "the ability to solve novel problems by using reasoning" (Kaufman, 2012, p. 119). Compared to measures that assess crystallized intelligence (i.e., "a knowledge-based ability that is highly based on education and acculturation"; Kaufman, 2012, p. 119), scores on the Raven's CPM depend less on academic knowledge and language (Naglieri & Ronning, 2000). However, it has to be noted that both fluid and crystallized intelligence contribute to the intellectual level of children. Therefore, Raven's CPM may differ in its relationships to VF performance compared to measures that assess crystallized intelligence (e.g., Wechsler Intelligence Scale for Children; Wechsler, 2003). For example, previous research has reported that parents influence fluid intelligence less strongly than crystallized intelligence (Rinderman, Flores-Mendoza, & Masur-Alves, 2010; Vista & Grantham, 2010). Future research into the relationship between VF performance and intelligence in young children should take these different intelligence measures into account.

While our measure of intelligence was found to have a positive influence on total word productivity, it did not affect the qualitative VF measures in our study. For the qualitative VF outcomes (i.e., mean cluster size, number of clusters, and number of switches), Raven's CPM score was deleted from the regression models since no predictive value of this score was found. Thus, the final models pertaining to the qualitative VF measures described in this study have not been controlled for intelligence. Nonetheless, among researchers examining executive functioning in children, no consensus has been reached yet as to whether intelligence (i.e., fluid or crystallized) should be taken into account when investigating these higher order functions in children (Arffa, 2007; Dennis et al., 2009; Mahone et al., 2002; Shanahan et al., 2006). Results of a study in children with attention deficit hyperactivity disorder indicated that the effect of the disorder on tasks of executive functioning (as is the VF task included in our study) is different for children with average and high IQ (Mahone et al., 2002). To ensure unbiased results, it might therefore be important to take IQ into account when examining performance on these kinds of tasks. On the other hand, other researchers have suggested that correcting for variation in IQ might lead to overcorrected results (Dennis et al., 2009).

Shanahan and colleagues (2006) suggested that analyses with an executive functioning measure, such as a VF task, as a dependent variable should be carried out with and without including intelligence to explore whether results would substantially differ. Therefore, we (re-)examined the influence of age, sex, LPE, and their interactions on the total number of correct words produced on the VF task, while leaving out Raven's CPM. In contrast to the result of the first analysis (described above), where the total number of words could be predicted by age and intelligence, our findings now revealed that, when intelligence was not taken into account, both age and LPE (especially low LPE) had a significant influence on total word productivity. The finding that lower parental education is connected to lower quantitative VF performance is in line with previous research (Ardila et al., 2005; Hurks et al., 2006; Klenberg et al., 2001). Furthermore, an interaction between sex and low LPE was found to be predictive for VF total word productivity. Taken together, results indicated that girls whose parents have a low LPE produce fewer correct words than all other children.

The finding that the outcome of the analysis changed by excluding intelligence shows that, as long as there is no consensus whether intelligence should be a covarying factor when examining cognitive development (and executive functioning more specifically) in children, researchers should carefully consider the exclusion or inclusion of intelligence in their analyses, and outcomes should be interpreted with caution. In our study, LPE correlated moderately with the scores on the intelligence measure. Also, the total variance explained by the model including Raven's CPM and by the model including LPE effects (without Raven's CPM) are highly comparable. These findings suggest that intelligence and LPE overlap in their influence on the total word productivity of young children and suggest that LPE might be merely a surrogate for the effects of intelligence on VF performance in young children. This association between Raven's CPM and LPE would (at least partially) explain why no significant effects of LPE were found for VF total word productivity when including Raven's CPM in the analysis. In other words, part of the variance of the total number of correct words that is associated with LPE might have been removed by entering intelligence in the model.

Next to the effect of LPE alone, an interactive effect of LPE and sex on the total word productivity emerged. A similar effect of LPE and sex, as discovered on the total

word productivity, was found on the number of switches. Girls with lower educated parents made fewer switches than other children. Taking the results together, it seems that low LPE only has influence on information retrieval processes and the use of switching strategies in girls. In line with previous studies (Brocki & Bohlin, 2004; Hurks et al., 2006; Lanting et al., 2009; Riva et al., 2000; Van der Elst et al., 2011), sex did not influence any of the VF outcome measures directly. Further research into the interaction between sex and LPE is necessary to determine the origin of this effect on total word productivity and number of switches.

It was surprising that switching (and the higher-order cognitive processes that are thought to underlie this VF outcome) was affected by LPE, since results from studies in older children have suggested that LPE mainly affects the total number of correctly generated words and the mean cluster size, which are thought to underlie lower-order cognitive functions such as lexico-semantic knowledge (Hurks et al., 2006, 2010). LPE can be seen as a proxy for other factors. For example, research in young children has found that there is a strong positive relationship between the level of education of the mother and the amount of time the mother interacts with the child (Hoff, 2003; Hoff & Tian, 2005). Furthermore, LPE has been shown to have a positive influence on the availability of reading material and behavior at home (Davis-Kean, 2005). The associations found in the present study between LPE and VF outcome measures might thus be mediated by these parent-child behaviors and therefore form an interesting starting point to further investigate the relationship between young children's VF performance and parent-related factors. Furthermore, since average LPE in our study was higher than the LPE estimates obtained for the general Dutch population, further examination of children's VF performance is required, specifically directed at children whose parents have lower LPE.

In conclusion, the present study gathered important information on the animal VF task for Dutch 4 to 6-year-old children, which helps to elucidate the cognitive development of young children. To our knowledge, this is the first study reporting data for children between the ages of 4 and 6 while using a variety of VF outcome measures (i.e., the total number of correct words, mean cluster size, number of clusters and number of switches). It is now relevant to investigate the clinical utility of this scoring method in the selected age group. By examining VF performance of young children suffering from a pathological condition, for example autism, attention deficit hyperactivity disorder or Down syndrome, knowledge can be acquired about the executive and associative processes important for VF performance in these specific groups of children.

REFERENCES

- Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336.
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology, 20*, 385-406.

- Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology*, 28, 539–560.
- Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology*, 22, 969–978.
- Baldo, J. V., Schwartz, S., Wilkins, D., & Dronkers, N. F. (2006). Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society*, 12, 896–900.
- Bowerman, B. L., & O'Connell, R. T. (1990). *Linear statistical models: An applied approach* (2nd ed.). Belmont, CA: Duxbury.
- Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, 26, 571–593.
- Centraal Bureau voor Statistiek [Statistics Netherlands]. (2012). *Stat-line*. Den Haag: Centraal Bureau voor Statistiek.
- Cherkes-Julkowski, M., Stolzenberg, J., & Segal, L. (1990). Prompted cognitive testing as diagnostic compensation for attention deficits: The Raven Standard Progressive Matrices and attention deficit disorder. *Learning Disabilities*, 2, 1–7.
- Davis-Kean, P. E. (2005). The influence of parent education and family income in child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19, 294–304.
- De Bie, S. E. (1987). *Standaardvragen 1987: Voorstellen voor uniformering van vraagstellingen naar de achtergrondkenmerken en interviews* [Standard questions 1987: Proposal for the uniformization of questions regarding background variables and interviews] (2nd ed.). Leiden, The Netherlands: Leiden University Press.
- Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, 15, 331–343.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378.
- Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders*, 38, 271–278.
- Hurks, P. P. M. (2012). Does instruction in semantic clustering and switching enhance verbal fluency in children? *The Clinical Neuropsychologist*, 26, 1019–1037.
- Hurks, P. P. M., Hendriksen, J. G. M., Vles, J. S. H., Kalff, A. C., Feron, F. J. M., & Kroes, M., ... & Jolles, J. (2004). Verbal fluency over time as a measure of automatic and controlled processing in children with ADHD. *Brain and Cognition*, 55, 535–544.
- Hurks, P. P. M., Schrans, D., Meijs, C., Wassenberg, R., Feron, F. J. M., & Jolles, J. (2010). Developmental changes in semantic verbal fluency: Analyses of word productivity as a function of time, clustering, and switching. *Child Neuropsychology*, 16, 366–387.
- Hurks, P. P. M., Vles, J. S. H., Hendriksen, J. G. M., Kalff, A. C., Feron, F. J. M., Kroes, M., ... & Jolles, J. (2006). Semantic category fluency versus initial letter fluency over 60 seconds as a measure of automatic and controlled processing in healthy school-aged children. *Journal of Clinical and Experimental Neuropsychology*, 28, 684–695.
- Kalff, A. C., Kroes, M., Vles, J. S. H., Bosma, H., Feron, F. J. M., Hendriksen, J. G. M., ... & Jolles, J. (2001). Factors affecting the relation between parental education as well as occupation and problem behaviour in Dutch 5- to 6-year-old children. *Social Psychiatry and Psychiatric Epidemiology*, 36, 324–331.
- Kaufman, J. C. (2012). Self estimates of general, crystallized, and fluid intelligences in an ethnically diverse population. *Learning and Individual Differences*, 22, 118–122.

- Kavé, G., Kligel, S., & Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *Journal of Clinical and Experimental Neuropsychology*, *30*, 349–359.
- Klenberg, L., Korkman, M., & Lahti-Nuutila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*, 407–428.
- Koren, R., Kofman, O., & Berger, A. (2005). Analysis of word clustering in verbal fluency of school-aged children. *Archives of Clinical Neuropsychology*, *20*, 1087–1104.
- Korkman, M., Lahit-Nuutila, P., Laasonen, M., Kemp, S. L., & Holdnack, J. (2013). Neurocognitive development in 5- to 16-year-old North American children: A cross-sectional study. *Child Neuropsychology*, *19*, 516–539.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). New York, NY: McGraw Hill.
- Lanting, S., Haugrud, N., & Crossley, M. (2009). The effect of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society*, *15*, 196–204.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.
- Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C., ... & Denckla, M. B. (2002). Effects of IQ on executive functions measures in children with ADHD. *Child Neuropsychology*, *8*, 52–65.
- Marquardt, D. W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association*, *75*, 87–91.
- Martens, R. (2012). *What counts? Cognitive development and arithmetical performance: The role of child- and environment-related factors* (Doctoral dissertation). Maastricht University, The Netherlands.
- Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Millis, S. R. (2003). Statistical practices: The seven deadly sins. *Child Neuropsychology*, *9*, 221–233.
- Mundfrom, D. J., Perret, J. J., Schaffer, J., Piccone, A., & Roozeboom, M. (2006). Bonferroni adjustments in tests for regression coefficients. *Multiple Linear Regression Viewpoints*, *32*, 1–6.
- Naglieri, J. A., & Ronning, M. E. (2000). Comparison of White, African American, Hispanic, and Asian children on the Naglieri Nonverbal Ability Test. *Psychological Assessment*, *12*, 328–334.
- Prigatano, G. P., Gray, J. A., & Lomay, V. T. (2008). Verbal (animal) fluency scores in age/grade appropriate minority children from low socioeconomic backgrounds. *Journal of the International Neuropsychological Society*, *14*, 143–147.
- Raven, J. C. (1958). *Guide to using the Coloured Progressive Matrices*. Oxford: H. K. Lewis & Co.
- Rinderman, H., Flores-Mendoza, C., & Mansur-Alves, M. (2010). Reciprocal effects between fluid and crystallized intelligence and their dependence on parents' socioeconomic status and education. *Learning and Individual Differences*, *20*, 544–548.
- Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language*, *71*, 267–284.
- Robert, P. H., Lafont, V., Medecin, I., Berthet, L., Thaubly, S., Baudu, C., & Darcourt, G. (1998). Clustering and switching strategies in verbal fluency tasks: Comparison between schizophrenics and healthy adults. *Journal of the International Neuropsychological Society*, *4*, 539–546.
- Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied Neuropsychology*, *12*, 190–201.

- Sauz on, H., Lestage, P., Raboutet, C., N’Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language, 89*, 192–202.
- Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit hyperactivity disorder? *Behavioural Brain Research, 130*, 3–28.
- Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., ... & De Fries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*, 585–602.
- Studenmund, A. H., & Cassidy, H. J. (1987). *Using econometrics: A practical guide*. Boston, MA: Little Brown.
- Tallberg, M., Carlsson, S., & Lieberman, M. (2011). Children’s word fluency strategies. *Scandinavian Journal of Psychology, 52*, 35–42.
- Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*, 370–378.
- Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138–146.
- Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499–504.
- UNESCO. (1997). *International Standard Classification of Education (ISCED)*. Paris: Author.
- Van der Elst, W., Hurks, P., Wassenberg, R., Meijs, C., & Jolles, J. (2011). Animal verbal fluency and design fluency in school-aged children: Effects of age, sex, and mean level of parental education, and regression-based normative data. *Journal of Clinical and Experimental Neuropsychology, 33*, 1005–1015.
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2006). Normative data for the animal, profession and letter m naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society, 12*, 80–89.
- Vista, A. D., & Grantham, T. C. (2010). Effects of parental education level on fluid intelligence of Phillipine public school students. *Journal of Psychoeducational Assessment, 28*, 236–248.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children – Fourth Edition. Administration and scoring manual*. San Antonio, TX: Harcourt Assessment, Inc.
- Weiss, E. M., Ragland, J. D., Bressinger, C. M., Bilker, W. B., Deisenhammer, E. A., & Delazer, M. (2006). Sex differences in clustering and switching in verbal fluency tasks. *Journal of the International Neuropsychological Society, 12*, 502–509.