Progress in Reading and Spelling of Dyslexic Children is Not Affected by Executive Functioning

Sietske A. E. Walda, Marjolijn van Weerdenburg, Maarten L. Wijnants, & Anna M. T. Bosman

Radboud University Nijmegen

Abstract

Although poor reading and spelling skills have been associated with weak skills of executive functioning (EF), its role in literacy is not undisputed. Because EF has different theoretical underpinnings, methods of analysis and of assessing EF, it has led to varying and often contrasting results in its effects in children with dyslexia. The present study has two goals. The first goal is to establish the relationship between a large number of EF tasks and reading and spelling skills in a large number of Dutch dyslexic children (n = 229). More interesting, however, is the second aim. To what extent do EF skills predict progress in reading and spelling in dyslexic children who attend a remediation program? The results reveal small, but significant relationships between EF and reading and spelling skills, but no relationships between EF and progress in reading and spelling. It is concluded that training EF skills is unlikely to enhance reading and spelling skills.

Keywords: Reading, writing, dyslexia, executive functions, working memory, longitudinal data analysis, special needs students

1. Introduction

Most authors view dyslexia as a developmental disorder that should be studied based on a cognitive causal model, in which deficiencies in phonological skills (e.g., letter-sound knowledge, phonological awareness, verbal short term memory, and rapid naming) are considered a main cause (Vellutino, Fletcher, Snowling, & Scanlon, 2004). However, several authors suggest that other deficits need to be considered. Deficiencies in sensory information processing (Ramus, 2003), motor and coordination (Ramus, 2003; Rochelle & Talcott, 2006), working memory (Gathercole, Alloway, Willis, & Adams, 2006), processing speed (Christopher et al., 2012), and attention (Franscheschini , Gori, Ruffino, Pedrolli, & Facoetti, 2012) are among a large list of examples associated with dyslexia.

Furthermore, several authors report that deficiencies in skills usually named Executive Functions or Executive Functioning (EF) often accompany dyslexia (Altemeier, Abbot, & Berninger, 2008; Berninger et al., 2006; De Weerdt, Desoete, & Roeyers, 2013; Jeffries & Everatt, 2004; Kegel &Bus, 2013; Menghini et al., 2010; Reiter, Oliver, & Lange, 2004). Although a uniform definition is lacking, EF is generally described as a set of cognitive processes that regulate non-automatic human behavior in a goal-directed and adaptive way (Banich, 2009; Best, Miller, & Jones, 2009; Hughes, 2001; Packwood, Hodgetts, & Tremblay, 2001; Miyake & Friedman, 2012), and is in many cases believed to be linked to the prefrontal cortex of the brain (Best et al., 2009; Miyake & Friedman, 2012). Although a phonological deficit is the most agreed upon cause of dyslexia, dyslexia is also associated with a plethora of other cognitive deficiencies, among which impairments in EF is one of the more recent possibilities.

1.1 Theoretical Background of Executive Functions

The first recorded example of a deficit in EF showed up in neuropsychological history when Phineas Gage in 1848 happened to have his brain pierced by an iron rod. Miraculously, he survived the brain damage when the rod was removed. However, his intimates reported changes in his personality. Twenty years later, John Harlow learned from the story of Gage and hypothesized the existence of prefrontal brain structures that allows one to plan and execute socially suitable behavior (Damasio, Grabowski, Frank, Galaburda, & Damasio, 1994). As more patients with frontal brain damage caused by accidents, medical conditions, war wounds, or frontal lobe surgery in mentally ill patients became apparent, more cases of injured rational decision-making and goal-directed behavior were reported (Feldman & Goodrich, 2001; Stuss, Gow, & Hetherington, 1991). Most patients tended to perform worse on tasks designed to assess prefrontal functioning (see for example, Gershberg & Shimanura, 1995; Goldstein, Bernhard, Fenwick, Burgess, & McNeill, 1993; Levine & Prueitt, 1989; Nauta, 1971). From these patients with frontal brain damage, it was hypothesized that lacking goal-directed behavior can be explained by deficiencies of the affected brain structures in the prefrontal cortex.

Nowadays, EF is no more solely studied in brain-damaged patients by neuropsychologists. EF has turned into a cognitive domain in psychology and psychiatry and is applied to patients with different diagnoses that imply some sort of “brain dysfunction”, like autism, ADHD, dyslexia, and dyscalculia. However, the fact that EF was first identified in brain-injured case studies leaves a few problems concerned with applying EF within other areas of expertise. In general, a great deal of the theoretical basis, validity, and assessment of EF is still indistinct according to recent review articles. First of all, a clear definition of EF is lacking and a number of different competing theories and models mutually coexist. As a consequence, there is no uniform approach to EF, which in turn leads to varying research results. Second, different research assumptions, questions, and methodologies are used causing a variety of different research results too. Third, whether EF form an overarching structure or take part in other cognitive constructs such as (fluid) intelligence, attention, and working memory is yet another issue in the domain (Banich, 2009; Hughes, 2011; Borkowski & Burke, 1996; Jurado & Rosselini, 2007; Lyon, 1996; Royall et al., 2002). We will discuss these issues in more depth here below.

To start with, no general definition of EF has been put forward. Different definitions highlight different attributes, such as changing strategies according to task demands (Borkowski & Burke, 1996), control mechanisms regulating cognitive behavior (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), cognitive functions that adapt human behavior to unfamiliar situations (Packwood, Hodgetts, & Tremblay, 2011), cognitive functions of the prefrontal cortex used for goal-directed behavior (Best et al., 2009), and the interplay of various cognitive functions needed for goal-directed behavior (i.e., attention and inhibition, task management, planning, updating working memory, and coding into working memory; Smith & Jonides, 1999). However, most authors agree that EF appears to be associated with cognitive functions necessary for performing goal-directed actions (Borkowski & Burke, 1996). These functions tend to develop into mature mental functions through childhood and adolescence (Jurado & Rosselini, 2007; Steinberg, 2004) and usually deteriorate with aging into seniorhood (Jurado & Rosselini, 2007; Treitz, Heyder, & Daum, 2007; but see Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2013 for a different view).

Second, there is some debate about the structure of EF. Different theories and models are in vogue with varying assumptions, questions, and methodologies (Borkowski & Burke, 1996), causing varying research results. In their review, Best et al. (2009) concluded that improvement in EF in childhood is associated with changes in the structure of EF (i.e., the relations among EF components). The structure of EF is of importance in the debate, because some authors argue that EF can be represented by one overarching mechanism responsible for navigating all goal-directed behavior (referred to as “the unity viewpoint”), whereas others state that EF can be understood as a coherent set of mental functions (referred to as “the nonunity viewpoint”). It is still not clear what functions are to be distinguished in the last case (Best et al., 2009).

Third, it is not clear whether EF is a unique construct or may take part in the concepts of (fluid) intelligence, attention, and working memory (Borkowski & Burke, 1996). Performances in EF tasks show high correlations with performances in tasks measuring fluid intelligence (skills that concern logical reasoning and problem solving in new situations, that are independent of previously learnt knowledge; Duncan, Emslie, Williams, Johnson, & Freer, 1996; Pennington & Ozonoff, 1996; Salthouse & Davis, 2006). This leads Salthouse and Davis (2006) to question the discriminant validity of the construct. One difficulty arises from the incongruent relation among intelligence, performance on EF-tasks, and frontal lobe damage. According to Sternberg (1985), executive processes form the basis of the general factor (g-factor) in tasks demanding intelligent behavior. Frontal lesions appear to have little influence on general measures of intelligence (WAIS IQ), but do affect measures of fluid intelligence (Duncan, Burgess, & Emslie, 1995). However, in healthy individuals general intelligence measures (WAIS-R IQ scores) are associated with EF-tasks (Obonsawin, Crawford, Page, Chalmers, & Cochrane, 2002), which resemble measures of fluid intelligence in many cases. Apart from intelligence, attention has been associated with EF at the conceptual, neuropsychological, neuroanatomical, and theoretical level (Barkley, 1996). Also working memory is considered part of EF, but due to measurement problems and ill definitions, a theoretical framework is far from complete. In particular, inhibition appears to play a linking role between working memory and EF. Also, several attentional systems overlap with EF and working memory within such a framework (Pennington, Bennetto, McAleer, & Roberts, 1996).

In short, it appears that EF cannot be strictly separated from other cognitive functions. (Parts of) fluid intelligence, attention, and working memory resemble EF, though none of these cognitive functions completely cover the concept. Likely, fluid intelligence, attention, working memory, and EF are not that clearly distinguished. Discussing them as well-defined cognitive constructs does not fit the actual situation.

Overall, the domain of EF has switched from neuropsychological practice to many adjacent work fields, including the field of learning disabilities. Due to the variety in definitions, multidimensionality of theoretical constructs, and overlap with other cognitive domains, the theoretical nature of EF remains unclear. In the next paragraph we will explain the state of affairs in research on EF.

1.2 Research concerning EF in general

Because of the large number of studies on EF, we concentrated on the most recent survey articles. The first survey pertaining to 46 studies on EF using confirmatory factor analysis (Royall et al., 2002) revealed no single overarching construct on which all EF tasks loaded. Instead, in most studies four factors of EF are found: “rule-discovery factor” (Wisconsin card sorting test), “working-memory factor”(California verbal learning test, Digit span, Tower of London), “attentional-control factor” (Continuous performance task, Digit cancellation), and “response-inhibition factor” (Digit span backward, Trial B, Stroop Task).

Packwood et al. (2011) adopted a different approach. A literature review on 60 studies on EF revealed 68 different terms describing EF. Using latent semantic analysis and hierarchical cluster analysis, the authors tried to extract the factors corresponding to the 68 terms, resulting in 18 factors. This, however, did not lead to a parsimonious solution. A more veracious approach would be to consider EF as a system responsible for guiding goal-directed behavior, independent of the specific behaviors necessary to perform EF tasks (Packwood et al., 2011).

In a third article, Miyake and Friedman (2012), solved the task-impurity problem by the “latent variable approach”, selecting “exemplar tasks” tapping into different EF and extracting common variance by performing confirmatory factor analysis. Three factors of EF were found by means of confirmatory factor analysis on results of the Wisconsin card sorting test, Tower of Hanoi, Random number generation, Operations span, and Dual task performance. These three factors were called "shifting" (most strongly related with Wisconsin card sorting test), "updating" (most strongly related with Randon number generation and Operation span), and "inhibition" (most strongly related with Random number generation and Tower of Hanoi). The three EF factors displayed a substantial amount of shared variance, suggesting a common underlying ability. However, the correlations between the three EF factors were far from perfect (ranging from r = .38 through r = .79), leaving room for diversity. They also correlated each in a different way with other tasks, such as IQ-tests and neuropsychological tests. The authors concluded that the three EF factors displayed both unity and diversity. In our view, this research forced the data from nine into three factors, with a partly overlap. Miyake and his colleagues found shared variance, which indicates unity. However, modest correlations between the EF factors indicate diversity. The statement that the EF factors display unity as well as diversity appears to be a contradiction.

The fourth and final survey article concerns a meta-analysis on 193 functional neuro-imaging studies (fMRI and PET). Using activation likelihood estimation, six EF's were found, namely “flexibility” (Task switching, Wisconsin card sorting test), “inhibition” (Antisaccades, Flanker task, GNT, Simon task, Stroop task), “working memory” (Complex calculation/PASAT, Delayed match to sample, N-back/AXCPT, Spatial span/sequence recall, Sternberg task), “initiation” (Word generation), “planning” (Tower maze test), and “vigilance” (Oddball discrimination) (Niendam et al., 2012). The results of this meta-analysis were interpreted as evidence in favour of a “common cognitive control network” that guides performance in all EF tasks, as revealed by increased activity in the associated brain structures during EF task performance. The authors also stress the probability of subdivisions in this common cognitive control network, all playing their own role in different EF's.

In sum, the review-studies concerning EF tasks revealed different meta-analytic approaches (confirmatory factor analysis, latent semantic analysis, hierarchical cluster analysis, structural equation modelling, and activation likelihood estimation) as well as approaches from different disciplines (neuropsychological studies and neuro-imaging studies). Also, many different methods for assessing EF were used. Limitations of tasks assessing EF, variability in meta-analytic approaches, various angles of view from concerned disciplines, and the use of many different assessment instruments have presumably contributed to the variety of different results produced by research on EF. In the next part, we will shed some light on limitations of assessment of EF by means of cognitive tasks, probably leading to the varying research results.

1.3 Limitations in assessment of EF by means of cognitive tasks

Tasks generally used for assessment of EF require several proficiencies, such as decision making, planning of actions, inhibition of automatic responses, switching from one problem solving strategy to another, focussing attention to important information, neglecting non-useful information, and so forth. On a conceptual level, assessment of EF by means of cognitive tasks is marked by severe limitations. These limitations partly lay in the way in which these tasks were historically designed. Patients with prefrontal brain damage were found to show curious behavior in rational decision-making, planning, social adaptation, concentration, among others, and performed worse on associated experimental tasks compared to control subjects. Therefore it was deduced that these experimental tasks measure EF (for examples see Gershberg & Shimanura, 1995; Goldstein, Bernhard, Fenwick, Burgess, & Mc Neil, 1993; Levine & Prueitt, 1989; Nauta, 1971), which has led experimenters to consider less proficient performances on these tasks as an indication of impaired EF.

Another limitation in the assessment of EF is the lack of a unequivocal definition and a commonly accepted theoretical framework that is domain-specific. For instance, Borkowski and Burke (1996) argued that EF cannot be directly measured in an individual, because each task measures concrete behavior, whereas information is desired about the ensemble of cognitive processes responsible for this behavior. Note that this problem concerns all psychological constructs (such as IQ and personality), not just EF.

Hence, assessing EF often results in severe task limitations. The first limitation of tasks designed to assess EF is the task-specificity problem, which reflects the uncertainty whether a task taps into the targeted EF and whether it taps into this EF alone (Miyake & Friedman, 2012; Packwood et al., 2011; Royall et al., 2002). A second limitation concerns task impurity, reflecting the problem that non-executive impairments could influence task performance as well as EF (Royall et al., 2002). Third, poor ecological validity is associated with those tasks, because the question remains whether poor test performance corresponds to poor abilities in real life (Hughes, 2011). Finally, many EF tasks require a reasonable level of (language) abilities, which cannot be demanded from children in the first grades of elementary school. Consider, for example, the Stroop color word interference test often used to assess EF. In this task, participants are shown words that name colours (e.g., GREEN, BLUE). These words, however, are not typed in regular in black font, but are coloured. Thus the word GREEN can be typed in green letters (congruent condition) or in yellow or blue letters (incongruent condition). Interference is measured by the participant’s slower color naming of printed words when confronted with incongruent color names compared to congruent colour names. Stroop interference is, however, associated with reading fluency (Protopapas, Archonti, & Skaloumbakas, 2007). Because most beginning readers do not meet this prerequisite, assessing EF in young students or in dyslexic students using this task may not be reliable. The conclusion seems justified that assessment of EF is complicated by a number of limitations that have not been resolved yet.

To sum up, it can be stated that research that merely uses EF tasks without additional information from questionnaires and observations, reveals no unitary concept of EF. Therefore, terms that are used in studies on EF cannot be reduced to an integrated account and share little similarities, which produces doubts about whether EF can be conceived as one construct or not. Based on the approach of Miyake et al. (2012) and Niendam et al. (2012), however, it is likely that EF tasks tap into skills that require an overarching control mechanism, although several factors of EF remain to be distinguished. This conclusion also fits the ideas of Packwood et al. (2011), namely, that EF pertains to a system responsible for guiding goal-directed behavior, independent of specific behaviors. Because of little consensus due to many different measures used, overlap between measures, and multidimensionality of constructs and tests, it appears impossible to provide a coherent view on EF.

1.4 Dyslexia, Improvement of Literacy skills, and EF

Many authors have reported on EF deficits in dyslexics (Brosnan et al., 2002; De Weerdt et al., 2013, Helland & Asbjørnsen, 2000; Jeffries & Everatt, 2004; Marzocchi et al., 2008; Menghini et al., 2010; Poljac et al., 2009; Reiter et al., 2004; Shanahan et al., 2006) and, more generally, on the role of EF in reading skills (Berninger et al., 2006; Foy & Man, 2013; Kegel & Bus, 2013). A meta-analysis of 48 studies on differences in EF between reading-impaired participants and controls showed that a number of studies found deficits in EF in dyslexic children, but others did not (Booth, Boyle & Kelly, 2010). It is postulated by the authors that this variety of outcomes at least partly originates from the many different tasks used for the assessment of EF. Impacts of several moderators on the relation between reading disorders and EF deficits were analysed, such as IQ-discrepancy criteria in defining reading disorder, whether it concerns a disorder of word reading or reading comprehension, age, gender, and the modality of the measurement task.

The results revealed that 91 out of 180 effect sizes were significant and indicated that the control groups tended to perform better than the reading-disorder groups. The coding subtest from the WISC-IV (Wechsler, 2004) appeared to discriminate best between control groups and reading-disorder groups. Studies that applied the IQ-discrepancy criterion, showed a smaller effect size than studies that did not apply the IQ-discrepancy criterion. Studies concerning a disorder of word reading on the one hand and a disorder of reading comprehension on the other did not differ significantly in effect sizes. Neither age nor gender influenced the effect sizes. However, response modality did. Children with reading disorders performed worse on EF tasks that required a verbal response than on tasks requiring a non-verbal response. Thus, the effect sizes seem to vary widely, because of different EF measurements used in the studies on EF, confirming the initial hypothesis of Booth et al. (2010). Also, the analyses revealed that the modality of the tasks used to measure EF, and whether an IQ-discrepancy criterion was used or not, influenced the magnitude of the effects. This lead the authors to the tentative conclusion that children with reading disorders have deficits in some areas of EF.

Despite a fair number of studies devoted to the relationship between EF and dyslexia, only two studies have considered the longitudinal role of EF in reading skills. The first study concerned beginning Chinese readers in Kindergarten 2 and 3 (Chung & McBride-Chang, 2011). The role of EF in reading acquisition was studied by controlling for other linguistic skills relevant for Chinese reading acquisition. The authors assessed "executive control" by an inhibitory-control task, which was based on the classic Stroop task but did not rely on proficiency of reading skills, as well as a working memory task, namely forward and backward digit span from the WISC-III (Wechsler, 1991). The linguistic skills included measures of phonological awareness, morphological awareness, and vocabulary. Reading acquisition was measured by a word-reading task. All tasks were administered in Kindergarten 2 (Time 1) and Kindergarten 3 (Time 2). Correlations of the linguistic skills and EF-tasks with word reading were moderate for Time 1 (ranging between r = .35 and r = .47) and Time 2 (ranging between r = .29 and r = .46). Also, correlations between measurements of linguistic skills and EF at Time 1 on the one hand and of word reading at Time 2 on the other, were moderate (ranging between r = .31 and r = .80).

Using regression analyses, the authors controlled for age. Other predictors were linguistic skills and EF. At Time 1, all predictors explained a significant part of the variance in word reading. At Time 2, only working memory did. Subsequently, word reading at Time 1 was added to the predictors in the regression model with word reading at Time 2 as dependent variable. The results showed that all other predictors turned out to be nonsignificant and the only significant predictor was word reading at Time 1. This suggests that linguistic skills and EF may be associated with reading skills, but cannot account for the improvement of word reading between Time 1 and Time 2. In another model, which included linguistic skills and EF at Time 1 as variables predicting word reading at Time 2, the final standardized beta coefficients of inhibitory control, working memory, and phonological awareness were significant. When in the last model word reading and age at Time 1 were added to the predictor variables, word reading and the two EF variables at Time 1 predicted significantly word reading at Time 2. However, F values of the EF variables were relatively small compared to the F values of word reading at Time 1. The authors stressed that the research was limited to Hong Kong children, and the results could only be applied to this group. They also concluded that, in future research more EF-tasks should be included in assessment of EF (Chung & McBride-Chang, 2011). In short, this study indicates that EF (inhibition and working memory) appears to be associated with reading skills in typically developing Chinese-speaking children, but future reading skills cannot be predicted by present EF.

The second study on the role of EF in literacy skills with some longitudinal basis concerns American children from Grades 1-4 (Cohort 1) and Grades 3-6 (Cohort 2), who completed four assessments of EF and literacy measures, each assessment separated by one year (Altemeier et al., 2008). Assessment of EF consisted of an inhibition task (i.e., a Stroop-task), an inhibition/switching task (switching quickly between naming the colour in which words are printed and reading words), and a rapid automatic switching task (switching between rapidly naming a high-frequency word and a double-digit number). Assessment of literacy skills consisted of word and text reading accuracy, word and pseudo-word reading rate, reading comprehension, written expression, and spelling.

From the results of regression analyses with three different predictor models (inhibition alone, entering inhibition followed by rapid automatic switching, and entering inhibition, and then rapid automatic switching, and then inhibition/switching), it was concluded that the measures of EF contributed to reading and writing level in typically developing children. However, results varied in the amount of explained variance for different grades and for different tasks. In comparison to the other literacy tasks, EF had weak predictive power for reading comprehension. Hierarchical linear modelling revealed that students who showed progress on all three EF's between Grade 1 and Grade 4, also showed better performance on literacy measurements in Grade 4, except for the effect of progress of inhibition on written expression. For dyslexic students, the same linear regression analyses were performed, indicating that EF contributes to reading and writing level too. However, associations in the dyslexic group were less strong than in the typically developing group. Also, the results indicate an effect of EF on reading and spelling progress only to a limited extent: Although multiple measurement moments are included in the study, no within-subjects differences in literacy skills at several measurement moments were evaluated. In this study, only between-subjects differences in literacy skills were evaluated.

The meta-analysis of studies on dyslexia and EF and the studies on the role of EF in literacy progress mentioned above show again that the interpretation of the results is hampered because of the variety of tasks used for assessment of EF. The studies also varied with respect to the spectrum of reading level, that is, reading-impaired children versus typically developing children.

1.5 The Present Study

The issues concerned with research on EF mentioned above will probably not be resolved in the near future. However, one can ask whether this is required to study EF in a meaningful way in relation to dyslexia. It may be more important to gain knowledge about the role of EF abilities in reading and spelling development in students with dyslexia. Establishing which specific EF tasks are associated with reading and spelling proficiencies appears ineffective, because results from previous studies depend on the EF task that is used for assessment. Insight in the relation between EF and progress on reading and spelling during dyslexia treatment, whether present or not, will lead to best practice training of literacy skills for students with reading and spelling difficulties and EF difficulties. Therefore, the aim of the present study was to assess to what extent EF measures relate to differences in a) reading and spelling level of dyslexic children and b) reading and spelling improvement during a remediation program.

2. Method

2.1 Participants

Participants were recruited from a Dutch organisation for assessment and remediation of learning disorders. Only clients who were eligible for repayment of diagnostics and/or treatment of dyslexia from health-care insurance companies were included. The children were referred and diagnosed during 2010 and 2011. Reading and spelling intervention was performed after the diagnostic phase. The final sample consisted of 229 participants, 73 girls and 156 boys, aged 7;1-11;1. Scores on a test for non-verbal reasoning (available for 203 participants) varied between T-score 30 and T-score 86, with a mean T-score of 53,5 (a standard score with a mean of 50 and a standard deviation of 10).

The participants are selected according to a strict procedure prescribed by the Dutch Care Insurance Act (Braams, 2008). The criteria are: a) participant receives education on a Dutch primary school; b) participant's results on word reading tests are within the 10^{th} percentile or participant's results on word reading tests are within the 16^{th} percentile and results on spelling tests are within the 10^{th} percentile; c) participant has received at least eight weeks of intensive and qualitative remediation on a one-to-one or small-group basis within the school without showing significant improvement on reading and spelling criteria above; (see criterion b); d) participant has no indication for another developmental disorder than dyslexia, which entitles for payment otherwise.

During the diagnostic process, it was evaluated whether criteria for inclusion of health care payment for reading and spelling intervention were met. The inclusion criteria were: 1) Intelligence quotient above 70; 2) participant's results on word reading tests within the 10^{th} percentile or participant's results on word reading tests within the 15^{th} percentile and results on spelling tests within the 10^{th} percentile, 3) low scores (within the 10^{th} percentile) on phonological processing tasks (two out of six measures) or, if not sufficed, low scores on memory tasks for phonological stimuli (e.g., sounds and syllables).

All participants met the criteria described in the Care Insurance Act. This resulted in the advantage of a relatively homogenous research group. Another advantage was that participants with reading and spelling difficulties who were subjected to poor education were excluded.

2.2 Procedure

Clients are registered for diagnosis and treatment by their parents. Parents were informed about the gathering of data for research and were assured that all data collection occurred anonymously. Prior to the reading and spelling intervention, parents also received information regarding the study. In case they wished not to join the study, they were excluded from the study. Less than 5% of selected participants refused to participate.

In all assessments, reading specialists served as test experimenters. The diagnostic procedure included measures of intelligence, EF (extensive details on these measures are presented in the materials' section), word and text reading, and spelling. The diagnostic sessions formed the pre-test of variables, followed by follow-up measures after approximately three (Follow-up 1), six (Follow-up 2), and nine months (Follow-up 3) of reading and spelling intervention (see Table 1).

Table 1

Design of Measured Variables

Pretest	Follow-up 1	Follow-up 2	Follow-up 3
Intelligence	n.a.	n.a.	n.a.
EF tasks	n.a.	n.a.	n.a.
Text reading	Text reading	Text reading	Text reading
word reading	word reading	Word reading	Word reading
Word spelling	Word spelling	Word spelling	Word spelling

Note. n.a. indicates that the task was not administered

2.3 Data Analysis

First, Pearson correlation coefficients between the performances on EF-computer tasks and reading and spelling measures were calculated to explore the extent to which EF relates to reading and spelling measures of dyslexic children.

Second, in order to investigate the extent to which EF measured by computer tasks relates to *improvement* of reading and spelling in dyslexic children, Pearson correlation coefficients were calculated and GLM repeated measures ANOVA's were run. The Pearson correlation coefficients were calculated between EF-computer task performances and *progress* on reading and spelling tests measures. To obtain a global indication on the progress that was made on reading and spelling, the differences between pre-test and follow-up 1, between follow-up 1 and 2, and between follow-up 2 and 3 were calculated. This type of analysis is not without discussion, but was chosen because the outcomes are easy to compare with the outcomes of the first step in the analyses (i.e., the correlations between EF and reading/spelling).

Finally, GLM repeated measures ANOVA's were carried out on the reading and spelling tests. The within-subjects factor contained four levels: Pretest, Follow-up1, Follow-up 2, and Follow-up 3; age was included as covariate because the dependent variables were not standardized. Consecutively, each of the EF measures was included as a covariate in a series of repeated measures to investigate the differences between measurements when controlling for that particular EF measure. The advantage of repeated measures ANOVA above correlations between EF measures and difference scores on reading and spelling, is that it controls for all potential sources of variability (i.e., treatment variability, within-subject variability, and random variability) and it reduces the variance of estimates of treatment-effects. Therefore, the power of a repeated measures ANOVA increases (Singer & Willett, 2003). Another advantage is that covariates can be easily included in the model.

Because of the clinical character of this study, missing data could not be avoided. Relatively many data were missing due to technical problems (data stored on various computers could not be retrieved), practical problems (because of lack of time within the given treatment sessions and drop-outs), and communication problems (additional instructions were not received by several experimenters, especially concerning Flanker-task letters). The assumption is warranted that missing data in the present study are random. Note, missing data were excluded from analysis, but participants with missing data were not excluded.

2.4 Materials

One test for nonverbal reasoning, three tests for working memory, 12 tests for EF, two tests for reading, and one spelling test were administered.

Intelligence. (Wechsler & Naglieri, 2008). Intelligence was assessed with the Matrix reasoning subtest of Wechsler nonverbal scale of ability (WNV-NL). In this test the participant had to evaluate how several shapes and geometric elements compare in a spatial or logical way. The goal was to choose from four to six options (depending on the item number of the task) the shape, which complemented the presented matrix. The test items were preceded by one demonstration item and three examples. Children older than six years started with item number 8. When item 8 and/or 9 were incorrect, the preceding items were presented until two consecutive items are correct. The test was discontinued after four errors within five consecutive items. Results were evaluated in terms of total amount of correct items.

Tower of London. The Tower of London (TOL, part of PEBL test battery: no 5, designed by Phillips, 1999; Mueller, 2010) is a computer task that consists of eight trials that increase in difficulty. The participant had to pile colored discs in three sections, building the same piles as in the example performing as few steps as possible. Also, only one disc at a time could be replaced and only the upper discs of the tiles could be replaced (see Figure A1 in the Appendix). Results were evaluated in terms of total number of steps necessary to complete all tasks, apart from the first one, and in terms of total reaction time necessary to complete these seven tasks.

Berg's card sorting task. The Berg's card sorting task (BCST, part of PEBL test battery, designed by Berg, 1948; Mueller, 2010) is a computer task that consists of four tiles of cards, illustrated with figures that vary in form (cross, circle, triangle, or star), number (one, two, three, or four), and color (red, blue, yellow, or green). Another card was presented, which had to be sorted on one of the piles, matching the principle of form, number or color of the printed figure(s). See Figure A2 in the Appendix. The participant had to find out which principle hds to be applied by the feedback ("correct" or "incorrect") given each time after having sorted a card on a pile. When correct, the same principle had to be applied to the next presented card. The computer changed the principle as soon as ten consecutive cards were sorted by the right principle, without notifying the participant. The new principle had to be found by the participant and applied to the next set of cards. The task was finished when all six sorting categories have been completed or when all 128 cards had been presented. Results were evaluated in terms of total count of correct responses and in terms of average reaction time.

Go-nogo task. The Go-nogo task is a computer task that was designed in analogy to Bezdjian, Baker, Lozano, and Raine (2009). In the present task, slightly different from the original, colors were used instead of letters, because of the possible interference of letter knowledge in children with dyslexia. The task for the participant was to respond to one stimulus and to withhold a response to the alternative stimulus. In this task, the participant was presented with yellow and blue circles. Within the first 50 trials, the participant was instructed to push the space bar as soon as possible when a yellow circle arose. When the computer shows a blue circle, nothing was to be done. After 50 presentations, the participant was instructed to do the reverse: The space bar had to be pushed when a blue circle was presented. Again 50 trials were presented. Prior to the experiment the participant received 15 probe trials prefaced by written as well as spoken instructions. The stimuli were shown for 500 ms, followed by a 2000 ms interval before the next stimulus was presented. Results were evaluated in terms of total count of nonresponses to target items (errors of omission, which is considered a measure of inattention, see Bezdjian, Baker, Lozano & Raine, 2009), total count of responses to non-target items, (errors of commission, which is considered a measure of impulsivity, see Halperin, Greenblatt & Young, 1991). Also, the average reaction time per trial was evaluated.

Flanker task arrows. The Flanker task ‘arrows’ is a computer task that was developed using E-prime, in analogy to Davelaar and Stevens (2009). The task consisted of 72 trials. After each 24 trials the task was paused. In each trial, five figures (->, <-, or -) were presented side by side. The participant was instructed to pay attention only to the center figure, and to press the correct button (->, or <-). The trials were congruent (target arrow flanked by same arrows), incongruent (target arrow flanked by opposite arrows), or neutral (target arrow flanked by dashes), each accounting for 24 trials (see Figure A3 in the Appendix). The experimental trials were preceded by 24 probe trials (8 congruent, 8 incongruent, and 8 neutral) and instructions were presented written as well as spoken. Stimuli were shown for 3000 ms, followed by a 2000 ms interval before the next stimulus was presented. Results were evaluated in terms of total number of correct responses to incongruent trials and average reaction time to incongruent trials, which are both considered to be measures of attentional control processes (Davelaar and Stevens, 2009).

Flanker task letters. The Flanker task ‘letters’ is a computer task that was developed using E-prime. The task resembled the above Flanker task arrows, however, in this task the capital letters “A” and “S” were being used as target items. Dashes were being used as neutral flankers (see Figure A4 in the Appendix). Results were evaluated in terms of total number of correct responses to incongruent trials and average reaction time to incongruent trials, which are both considered to be measures of attentional control processes (Davelaar and Stevens, 2009).

Digit recall. Digit recall was assessed using a subtest of the Working memory test battery for children (WMTB-C; Pickering & Gathercole, 2001). The participant wss asked to repeat a sequence of digits read aloud by the experimenter, in the same order. After two practice items, sequences of three digits were to be repeated. When less than four of six presented sequences are repeated correctly, the participant proceeded to the test items with two digits. When at least four of six presented sequences were repeated correctly, the reading specialist proceeded to the test items with one more digit, until less than four sequences were repeated correctly.

Backward recall. Backward recall was assed using a subtest of the Working memory test battery for children (WMTB-C; Pickering & Gathercole, 2001). The participant was asked to repeat a sequence of digits read aloud by the reading specialist in reverse order. The procedure of test administration resembled that of digit recall.

Blockrecall. Digit recall was assed using a subtest of the Working memory test battery for children (WMTB-C; Pickering & Gathercole, 2001). The participant had to imitate the sequence of blocks positioned on a board pointed out by the reading specialist. The procedure of test administration resembled that of digit recall.

Naming speed. Naming speed was assessed by a subtest of Differential Diagnostics of Dyslexia [benoemsnelheid, Differentiaal Diagnostiek van Dyslexie, 3DM] (Blomert & Vaessen, 2009). This computer task consisted of three subtasks in which the participant is asked to name as quickly as possible letters, digits, and pictures respectively. Stimuli were presented on the computer screen. Each task was presented twice, resulting in reaction times for the total of two letter tasks, two digit tasks, and two picture tasks.

Coding. Coding was assessed using a subtest of the Wechsler Intelligence Scale for Children (WISC-III NL) (Wechsler, 2005). At the top of this paper-and-pencil test, the numbers 1-9 were presented, each corresponding to another symbol. Subsequently, the participant was presented with random numbers 1-9, and each time had to print the corresponding symbol as quickly as possible, analogous to the model printed on top of the page. Results were evaluated in terms of total number of correctly printed symbols within two minutes.

Score double task. The Score double task is a subtest of the Test of Everyday Attention for Children (TEA-Ch) (Manly, Robertson, Anderson, & Nimmo-Smith, 1999/2004). In this task the participant was asked to listen to ten radio news items on an audio CD and to remember from each news item the animal that was named, as well as the number of special sounds, which they were told was a sound of a computer game. After each item the participant was asked to name the animal and the number of sounds. Results were evaluated in terms of total number of correctly named animals and the number of correctly remembered sets of sounds.

Opposite worlds. The Opposite worlds task is a subtest of the Test of Everyday Attention for Children (TEA-Ch) (Manly et al., 1999/2004). In this task the participant was presented a path of squares in which “1” or “2” is printed. In the real-world condition, the participant had to point the squares one at a time while naming the digits as quickly as possible. In the opposite-world condition, the participant had to say “1” when presented with a 2, and “2” when presented with a 1. The participant had to complete four paths: Two of each condition. The experimenter registered reaction times using a stopwatch. Reaction times of the four paths were summed resulting in an opposite worlds inhibition score.

Word reading 1-3. Word-reading was assessed with the Three Minute Test [Drie Minuten Toets, DMT] (Jongen & Krom, 2009). This test consisted of three different cards containing single words. Card 1 contained VC- and CVC-words (C is a consonant and V is a vowel), Card 2 contained CCV, CCVC, CVCC, CCVCC, CCCVC and CVCCC words, and Card 3 contained multisyllabic words. The words had to be read aloud as fast as possible. Test score for each card was the number of words read accurately within one minute.

Word reading 4. The fourth word-reading task is called the One-Minute Test [Eén Minuut-Test, EMT] (Brus & Voeten, 1972). This test consisted of words listed on a sheet, in ascending order of difficulty. The words had to be read aloud as fast as possible. Test score is the number of words read accurately within one minute.

Text reading. Text reading was assessed by one text of AVI test package [AVI toets pakket] (Visser, Van Laarhoven, & Ter Beek, 1996). The test contains several test cards with text in ascending order of difficulty. All participants read text card 3A, which they had to read as accurately and as fast as possible. Test scores were reading time and the number of errors.

Word spelling. Spelling of words was assessed with the PI test [PI-dictee] (Geelhoed & Reitsma, 2004). This is a writing-down-to-dictation task of a list of single words. The entire test contained 9 lists of 15 words increasing in difficulty, covering spelling categories from Grades 1 to 6. The test was discontinued when fewer than 8 words were spelled correctly in a list. Test score was the number of words spelled correctly.

3. Results

3.1 Descriptives

Table 2 shows levels of reading and spelling of the participants at Pretest and Follow-up 1-3 indicated by categories of percentiles. These categories correspond to A, B, C, D, and E-scores often used in Dutch education, with A corresponding to the highest category and E corresponding to the lowest. It is noteworthy that, according to the dyslexia protocol approved by Dutch health care insurances, one of the purposes of reading and spelling intervention is to minimize E-scores on reading and spelling (Nationaal Referentiepunt Dyslexie, 2013). Table 2 shows 88.2% of participants who started intervention, had an E-score on spelling. After roughly nine months of reading and spelling intervention, 32.9% of the participants scored E. In most cases, reading and spelling intervention was continued following these nine months. Table 2 also shows results on reading words 1-3, revealing progress on word reading for most participants was less than progress on word spelling. These results show that many participants had a persistent reading disorder. However, progress of reading and spelling is compared to peers. Participants that scored E after nine months of reading and spelling intervention, in most cases did show individual progress on reading and spelling measures compared to the results at the Pretest. Participants who scored D, C, B, or A after nine months of reading and spelling intervention even outpaced regular speed of reading and spelling progress of peers. Moreover, clinical experience shows that dyslexic participants show more progress in reading text compared to reading words. According to the dyslexia protocol, a subsequent goal of the reading and spelling intervention is to facilitate functional literacy by teaching strategies to cope with reading and spelling difficulties (Nationaal Referentiepunt Dyslexie, 2013). However, results of reading and spelling intervention on text reading are not available because of difficulties in comparing performance by standardized norm scores. In interpreting this results, one has to take into account the Matthew effect that applies to reading development (the fact that skilled readers tend to have a cumulative advantage in their reading progress compared to less-skilled readers, see Stanovich, 1986). The results of the participants in comparison to their peers would likely have been far worse if they had not received intensive reading and spelling intervention.

Table 2

Counts and Percentages of Levels of Reading and Spelling

		Pretest		FU1[a]		FU2[a]		FU3[a]	
	percentile	*N*	percentage	*n*	percentage	*n*	percentage	*N*	percentage
Word spelling	75-100	0	0.0	1	0.4	5	2.2	5	2.2
	50-75	0	0.0	8	3.5	7	3.1	18	7.9
	25-50	9	3.9	26	11.4	42	18.4	48	21.1
	10-25	16	7.0	27	11.8	51	22.4	51	22.4
	0-10	201	88.2	165	72.4	118	51.8	75	32.9
	missing	2	0.9	1	0.4	5	2.2	31	13.6
Word reading 1	75-100	1	0.4	2	0.9	1	.04	2	0.9
	50-75	1	0.4	2	0.9	3	1.3	8	3.5
	25-50	8	3.5	29	12.7	26	11.4	23	10.1
	10-25	26	11.4	44	19.3	51	22.4	46	20.2
	0-10	147	64.5	139	61.0	131	57.5	110	48.2
	missing	45	19.7	12	5.3	14	0.9	39	17.1
Word reading 2	75-100	0	0.0	1	0.4	2	0.9	2	0.9
	50-75	0	0.0	3	1.3	3	1.3	5	2.2
	25-50	11	4.8	26	11.4	21	9.2	23	10.1
	10-25	29	12.7	48	21.1	53	23.2	45	19.7
	0-10	142	62.3	137	60.1	134	58.8	114	50.0
	missing	46	20.2	13	5.7	15	6.6	39	17.1
Word reading 3	75-100	0	0.0	1	0.4	0	0.0	2	0.9
	50-75	0	0.0	2	0.9	2	0.9	1	0.4
	25-50	3	1.3	11	4.8	13	5.7	21	9.2
	10-25	23	10.1	48	21.1	56	24.6	41	18.0
	0-10	144	63.2	146	64.0	136	59.6	122	53.5
	missing	58	25.4	20	8.8	21	9.2	41	18.0

[a] FU means Follow-up

Note that Table 2 shows varying amounts of missing values. The causes of most missing values were threefold. First, at the Pretest the word reading 1, 2, and 3 measures were not part of the standard diagnostic battery, so in a number of cases these tests were not administered. It was assumed that these missing values were at random. Second, in several cases word reading 3 turned out to be too difficult because of increasing difficulty of the word reading tests. So it can be expected that word reading 3 scores miss in cases of the most severe reading difficulties. Third, when reading and spelling intervention goals were achieved, the intervention was stopped. Therefore, some missing results at Follow-up 2 and many missing results at Follow-up 3 can be expected to be relatively high because intervention goals had been already met.

3.2 Correlational Analysis

Table 3 shows Pearson correlations between the reading and spelling tests (at Pretest and Follow-ups 1, 2, and 3) and the EF measures (at Pretest), resulting in a correlation table with 504 cells. Analysis of the correlations reveals that 388, or 80,0% of these correlation coefficients reach a significance level of $p < .05$.

Table 4 shows Pearson correlations between progress on reading and spelling tests and the EF-measures at Pretest. Progress on reading and spelling tests were computed between Pretest and Follow-up 1, Follow-up 1 and Follow-up 2 and Follow-up 2 and Follow-up 3. The correlation table counts 378 cells. Analysis of the correlations reveals that 31, or 8,2% of the correlation coefficients reach a significance level of $p < .05$.

Table 3

Correlations between Executive-Functions Tasks and Reading and Spelling Level

		Word reading 1				Word reading 2				Word reading 3				Word reading 4				Text reading			Word spelling		
	Pretest	Pretest	FU 1[a]	FU 2[a]	FU 3[a]	Pretest	FU 1[a]	FU 2[a]	FU3[a]	FU 1[a]	FU 2[a]	FU 3[a]	Pretest	FU 1[a]	FU 2[a]	FU 3[a]	Pretest	FU 1[a]	FU 2[a]	FU 3[a]	Pretest	FU 1[a]	FU 2[a]	FU 3[a]
Tower of London																								
Number of steps	-.15*	.19*	.12	.08	.19*	-.15*	-.17*	-.13	-.15*	-.17*	.18*	-.15*	-.10	-.08	-.05	-.05	-.09	-.06	-.05	-.02	-.15	-.13	-.15	-.10
Total reaction time	-.40**	.35*	.37**	.22**	.35**	-.39**	-.41**	-.39**	-.39**	-.32**	.24*	-.40**	-.34**	-.30**	-.30**	-.21*	-.35**	-.32**	-.29**	-.22**	-.37**	-.37**	-.35**	-.27**
Berg's card sorting test																								
Number correct	.24**	-.25**	-.26**	-.25**	-.25**	.25*	.26**	.31**	.25**	.28**	-.14	.24**	.34**	.17*	.21**	.25**	.30*	.25**	.24**	.25**	.32**	.27**	.28**	.29**
Mean reaction time per trial	-.43**	.45**	.33**	.37**	.45**	-.42**	-.41**	-.40**	-.42**	-.33**	.28**	-.43**	-.09	-.01	-.06	.01	.05	-.07	-.03	.07	-.07	-.09	-.07	-.03
Go-nogo test																								
Number of omission errors	-.25**	.16	.27**	.25**	.16	-.26**	-.28**	-.31**	-.26**	-.32**	.16	-.25**	-.26**	-.27**	-.26**	-.23**	-.28**	-.27**	-.24**	-.17*	-.28**	-.31**	-.30**	-.23**
Number of commission errors	-.16*	.19*	.08	.13	.19*	-.20**	-.26**	-.20**	-.20**	.17*	.21*	-.16*	-.17*	-.20**	-.23**	-.16*	-.15	-.19**	-.15*	-.09	-.15	-.20**	-.15	-.23
Mean reaction time per trial	-.13	.06	.06	.03	.06	-.18*	-.12	-.13	-.18*	.16*	-.08	-.13	-.14	-.07	-.06	-.01	-.19**	-.12	-.09	.00	-.21**	-.15*	-.16*	-.07
Flanker test letters																								
Number of correct responses	-.35**	.36**	.21*	.16	.36**	-.21*	-.25**	-.20*	-.21*	.31**	-.00	-.35**	.06	-.03	-.06	.05	-.01	.02	-.06	.02	-.08	-.02	-.05	.01
Mean reaction time per trial	-.08	.16	-.00	-.04	.16	-.02	-.02	-.04	-.02	-.03	-.06	-.08	-.29**	-.28**	-.24**	-.23*	-.35**	-.31**	-.26**	-.24*	.34**	-.29**	-.28**	-.17
Flanker test arrows																								
Number of correct responses	-.29**	.21*	-.20**	.18*	.21*	-.21	-.26**	.24**	-.21**	-.21**	.05	-.23**	.27**	-.32**	.30**	.34**	.28**	.29**	.27**	.30**	.17*	-.25**	.26**	.26**
Mean reaction time per trial	.21**	.19*	.30**	-.35**	-.19*	.17*	.24**	.25**	.172*	.20**	-.24**	.21**	-.26**	-.24**	-.27**	-.24**	-.30**	-.24**	-.27**	-.21*	-.32**	-.27**	-.30**	-.19*
Digit recall	.15*	.-.13	-.11	-.11	-.06	.20*	.25**	.22**	.23**	.16*	.18*	.17*	.22**	.12	-.13	.14	.20**	.13	.14	.11	.19*	.16*	.16*	.16*
Backward digit recall	.26**	-.19*	-.11	-.11	-.13	.27*	.28**	.23**	.15	.21**	.23**	.13	.23**	.21**	.20**	.10	.22**	.23**	.18*	.15	.26*	.23*	.22**	.20*
Block recall	.25**	-.32**	-.23**	-.29**	-.22*	.28*	.31**	.34**	.21**	.19**	.28**	.17*	.29**	.22**	.27**	.12	.23**	.25**	.29**	.15	.21*	.23*	.25**	.17*
3DM naming speed letters	-.08	.15	-.16*	.17*	.19*	-.02	-.06	-.04	.05	-.12	-.09	-.13	-.09	-.09	-.10	-.11	-.09	-.09	-.09	-.13	.10	-.11	-.13	-.16*
3DM naming speed digits	-.43**	.34**	.31**	.29**	.27**	.30**	-.35**	.32**	-.32**	-.40**	-.43**	-.29**	-.39**	-.42**	-.63**	-.26**	-.41**	-.43**	-.56**	-.26**	-.40**	-.37**	-.53**	-.25**
3DM naming speed pictures	-.38**	.45**	.38**	.31**	.39**	-.39**	-.33**	.33**	-.25**	.34**	-.32**	-.25**	-.42**	-.37**	-.41**	-.32**	-.43**	-.38**	-.38**	-.28**	-.40**	-.36**	-.35**	-.22**
WISC-III coding	.16*	-.37**	-.13	-.07	-.14	.26**	.26**	.17:	.14	.12	.11	.09	.21**	.08	.08	.05	.28**	.11	.10	.09	.30**	.25**	.20*	.08
TEA-Ch score double task number of animals	.22**	-.18*	-.24**	-.32**	-.25**	.16*	.21**	.21**	.17*	.20**	.21**	.16*	.21**	.19*	.21**	.16	.24**	.29*	.24**	.20**	.18*	.31**	.23**	.21*
TEA-Ch score double task number of sounds	.24**	-.22*	.22**	-.22**	-.12	.23**	.22**	.19**	.27**	.19**	.18*	.19*	.24**	.10	.20**	.17*	.28**	.18*	.20**	.17*	.22**	.17*	.18*	.13
TEA-Ch opposite worlds inhibition score	-.47**	.45**	.50**	.45**	-.46**	-.35**	-.42**	-.43**	-.40**	.43**	-.43**	-.39**	-.49**	-.47**	-.49**	-.43**	-.43**	-.46**	-.42**	-.38**	-.43**	-.48**	-.44**	-.35**

Note. * p < .05 ** p < .01. Yellow marked *p*-values indicate significant correlations on at least p < .05 level. [a] FU means Follow-up

Table 4

Correlations between Executive-Functions Tasks and Reading and Spelling Progress Throughout the Intervention

	Word reading 1			Word reading 2			Word reading 3			Word reading 4			Text reading			Word spelling		
	P-FU1[a]	FU1-2[a]	P-FU1[a]	FU1-2[a]	FU2-3[a]	P-FU1[a]	FU1-2[a]	FU2-3[a]	FU2-3[a]	P-FU1[a]	FU1-2[a]	FU2-3[a]	P-FU1[a]	FU1-2[a]	FU2-3[a]	P-FU1[a]	FU1-2[a]	FU2-3[a]
Tower of London																		
number of steps	-.01	-.01	-.10	.00	-.02	-.05	.03	.01	.15	-.00	.04	-.07	-.03	-.00	.02	-.03	-.11	.11
Total reaction time	.12	.14	-.16	-.23**	-.10	-.02	-.03	.01	.18*	.02	.03	.13	-.03	.03	.05	-.08	.00	.13
Berg's card sorting test																		
number correct	.15*	.07	.17	.14	.02	.03	.13	-.04	.02	-.04	.15*	.12	.11	.04	.03	.22**	.07	.04
Mean reaction time per trial	.15*	.03	-.43**	-.21*	-.18*	.11	-.05	-.01	-.01	.23**	.03	.15	.05	.04	.04	.03	-.02	.01
Go-nogo test																		
Number of omission errors	-.10	-.01	-.05	-.03	-.13	.02	-.01	.06	.10	-.00	.03	.04	.01	.06	.16	-.09	-.02	.06
Number of commission errors	.01	.04	-.17	-.05	-.00	-.01	.04	-.12	-.02	-.01	-.11	.07	-.02	.07	-.03	-.09	.10	.02
Mean reaction time per trial	-.05	-.04	.04	-.01	-.00	.08	.01	.01	.09	.08	.04	-.02	.03	.10	.12	-.03	.04	.02
Flanker test letters																		
Number of correct responses	.02	.04	-.35**	-.18	.09	-.01	.06	-.10	.02	.06	.08	-.06	.05	.10	-.14	.08	-.00	.05
Mean reaction time per trial	.06	.04	-.12	.01	.02	-.04	-06	.24*	.22*	.13	-.16	.23*	-.02	-.13	.15	.07	-.04	.03
Flanker test arrows																		
Number of correct responses	-.02	-.11	-.10	-.09	.03	-.04	-.01	-.11	.04	-.05	-.03	-.02	-.02	-.06	.00	-.10	-.08	.16
Mean reaction time per trial	.03	-.03	.11	.13	.23**	.10	.05	.06	.08	.09	-.04	.04	.04	-.04	-.00	.16	.01	-.10
Digit recall	.01	.05	.08	.05	.08	.08	-.07	.08	-.00	-.12	.00	.05	-.08	.02	-.03	.02	-.02	.07
Backward digit recall	-.06	.06	.14	.06	-.04	-.03	-.07	.02	-.10	-.01	-.03	-.17*	.03	-.13	.10	.01	-.03	.06
Block recall	-.04	-.18*	.14	.18*	.18*	.05	.02	.04	-.01	-.05	.07	-.08	.06	.06	-.08	.08	-.04	-.03
3DM naming speed letters	-.07	.06	-.05	.06	-.02	-.08	.04	-.02	-.08	-.02	-.03	.00	-.03	-.02	-.08	-.07	-.07	-.08
3DM naming speed digits	.06	-.16*	-.16	-.17*	.01	-.05	.04	.28**	.04	-.03	-.15*	.12	.01	-.03	-.09	.03	-.01	.07
3DM naming speed pictures	.06	-.02	-.24**	-.02	-.02	.03	-.07	-.03	-.02	.08	-.12	.03	.04	-.03	-.02	-.07	-.03	.07
WISC-III coding	-.03	-.05	.24*	-.05	-.00	-.05	-.06	.08	.08	-.18*	.08	-.00	-.17*	.08	.10	-.05	-.05	-.02
TEA-Ch score double task number of animals	-.06	.08	.06	.08	.28**	.06	.05	.08	-.05	-.05	.11	-.04	.08	-.06	.00	.10	-.11	.10
TEA-Ch score double task number of sounds	-.07	.02	.15	.02	.12	-.01	-.04	.18*	-.03	-.16*	.22*	.02	-.07	.08	-.04	.01	-.06	-.05
TEA-Ch opposite worlds inhibition score	.06	-.08	-.27**	-.08	-.15	-.09	-.06	-.04	.05	.07	-.12	.02	-.07	-.00	-.03	-.11	.07	.04

Note. * $p < .05$ ** $p < .01$. Yellow marked *p*-values indicate significant correlations on at least p < .05 level. [a] FU means Follow-up

3.3 Development of Reading and Spelling

Results of repeated measures on reading and spelling tests at Pretest, Follow-up 1, Follow-up 2, and Follow-up 3 with age as a covariate are shown in Table 5. Age was added as a covariate because the reading and spelling measures were not standardized. Because the focus of this study is on effects of EF on reading and spelling progress, interaction effects with age were not reported. All reading and spelling measures showed significant progress during the first three months of reading and spelling intervention. The participants showed faster reading and more accurate spelling, indicated by increasing scores on word reading and word spelling. Scores on text reading indicated faster reading as well, because the children required less time to complete the text. Within the second three-month period of reading and spelling intervention, children only progressed significantly on text reading. During the last months of reading and spelling intervention, children progressed only significantly on word reading 1.

Table 6 shows whether interaction effects with the EF variables were significant. In most cases, interaction effects did not reach significance: only 21 *p*-values (5.6%) did and this is a number that is expected just by chance. The conclusion that performance on EF measures does not affect progress on reading and spelling measures during reading and spelling intervention is warranted.

Table 5

Effect of Dyslexia Treatment on Reading and Spelling Measures at Pretest and Follow-Ups 1 through 3 with Age as a Covariate

		N	Pretest	FU1[d]	FU2[d]	FU3[d]	Pretest-FU1[d]		FU1-FU2[d]		FU2-FU3[d]	
Measure			*Mean (SD)*	*Mean (SD)*	*Mean (SD)*	*Mean (SD)*	*F*	*p*	*F*	*p*	*F*	*p*
Word reading[a]	1	166	45.1 (18.1)	57.0 (18.6)	63.1 (19.3)	66.9 (18.2)	14.57	.00**	.06	.81	4.08	.05*
	2	136	32.7 (17.9)	45.1 (19.1)	50.3 (20.3)	54.7 (19.8)	22.17	.00**	.94	.34	3.19	.08
	3	135	24.1 (14.5)	34.2 (16.2)	39.7 (16.9)	43.4 (16.8)	20.19	.00**	.10	.75	.20	.66
	4	125	27.5 (11.4)	34.9 (11.6)	39.2 (12.5)	41.0 (11.4)	6.12	.02*	.01	.92	1.89	.17
Text reading [b]		83	134.2 (71.0)	98.3 (50.3)	81.4 (36.0)	73.8 (32.3)	28.32	.00**	16.54	.00**	.94	.34
Word spelling [c]		170	41.0 (21.0)	63.7 (20.0)	75.4 (20.6)	85.4 (19.7)	12.10	.00**	.07	.80	3.79	.05

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

[a] Evaluated in terms of total count of correct words. [b] Evaluated in terms of total time required to read the text. [c] Evaluated in terms of total count of correct words. [d] FU means Follow-up

3.4 The Role of EF in Development of Reading and Spelling

Introducing each EF measure separately as a covariate in the repeated measures analyses for reading and spelling measures resulted in 378 separate analyses. Table 6 shows that most effects remained unchanged when doing this. This indicates that in the majority of cases, performance on the EF-measures does not affect progress on reading and spelling measures during reading and spelling intervention. However, 50 analyses (or 13.2%) revealed that training effects were different from those in Table 5. Of these *p*-values, 36 (or 9,5%) turned from significant into nonsignificant values. In these cases, the significant effect of reading and spelling intervention on reading and spelling progress found in Table 5 can be explained by the concerned EF-skill. On the other hand, 14 of these *p*-values (or 3,7%) turned from nonsignificant to significant values. In these cases, the initial nonsignificant effect of reading and spelling intervention on reading and spelling progress turned out to be related to differences in the concerned EF-skills. Apparently, in these analyses, students with varying EF skills differ in the amount of progress during reading and spelling intervention.

Table 6

Progress on Reading and Spelling Measures During Reading and Spelling Intervention, at Pretest and Follow-ups 1-3, with Age and EF Measures as Covariates

Measure		*N*	Pretest-FU1[d]				FU1-FU2[d]				FU2-FU3[d]			
			F	*p*	*int age*[c]	*int EF*	*F*	*p*	*int age*[c]	*int EF*	*F*	*p*	*int age*[c]	*int EF*
Tower of London number of steps														
Word reading[a]	1	146	8.56	.00**	n	n	0.00	.97	n	n	1.53	.22	n	n
	2	115	14.99	.00**	y	n	0.13	.72	n	n	5.25	.02*	n	n
	3	114	13.89	.00**	y	n	0.48	.49	n	n	0.13	.72	n	n
	4	105	5.46	.02*	n	n	0.01	.94	n	n	1.63	.21	n	n
Text reading [b]		80	19.11	.00**	y	n	13.99	.00**	y	n	0.44	.51	n	n
Word spelling [c]		149	13.22	.00**	n	n	0.03	.86	n	n	2.46	.12	n	n
Tower of London total reaction time														
Word reading[a]	1	146	3.18	.08	n	n	0.81	.37	n	n	1.22	.27	n	n
	2	115	12.68	.00**	y	n	0.13	.72	n	n	2.00	.16	n	n
	3	114	17.30	.00**	y	n	0.06	.81	n	n	0.35	.56	n	n
	4	105	3.83	.05	n	n	8.08	.61	n	n	28.47	.35	n	n
Text reading [b]		80	17.06	.00**	y	n	7.98	.01**	y	n	0.03	.86	n	n
Word spelling [c]		149	10.08	.00**	n	n	0.10	.75	n	n	2.57	.11	n	n
Berg’s card sorting test number of correct responses														
Word reading[a]	1	122	16.05	.00**	y	n	0.22	.64	n	y	2.28	.13	n	n
	2	122	11.89	.00**	y	n	0.22	.64	n	n	0.04	.85	n	n
	3	113	1.23	.27	n	y	0.05	.82	n	n	0.60	.44	n	n
	4	151	9.27	.00**	y	y	0.40	.53	n	n	2.06	.15	n	n
Text reading [b]		76	20.64	.00**	y	n	12.83	.00**	y	n	1.94	.17	n	n
Word spelling [c]		153	10.18	.00**	n	n	16.78	.66	n	n	3.51	.06	n	n
Berg’s card sorting test mean reaction time per trial														
Word reading[a]	1	124	7.09	.01**	y	n	0.40	.53	n	n	0.45	.50	n	n

	2	123	14.48	.00**	y	n	0.01	.93	n	n	0.15	.70	n	n
	3	114	2.24	.14	n	n	0.82	.37	n	n	0.07	.79	n	n
	4	152	5.54	.02*	n	n	0.00	.96	n	n	2.65	.11	n	n
Text reading [b]		77	6.38	.01*	y	y	5.75	.02*	y	n	0.31	.58	n	n
Word spelling [c]		154	5.65	.02*	n	n	0.08	.79	n	n	2.95	.09	n	n
Go-nogo test number of omission errors														
Word reading[a]	1	122	21.64	.00**	y	n	0.46	.50	n	n	3.11	.08	n	n
	2	121	24.22	.00**	y	n	0.12	.73	n	n	0.51	.48	n	y
	3	112	5.82	.02*	n	n	0.01	.94	n	n	1.23	.27	n	n
	4	149	14.16	.00**	y	n	0.01	.91	n	n	2.23	.14	n	n
Text reading [b]		78	36.76	.00**	y	n	12.66	.00**	y	n	0.07	.80	n	n
Word spelling [c]		153	10.72	.00**	n	n	0.00	.99	n	n	2.56	.11	n	n
Go-nogo test number of comission errors														
Word reading[a]	1	122	20.96	.00**	y	n	2.16	.14	n	y	2.68	.11	n	n
	2	121	24.39	.00**	y	n	0.03	.87	n	n	0.06	.81	n	n
	3	112	7.91	.01**	n	n	0.31	.58	n	n	2.77	.10	n	n
	4	149	9.67	.00**	n	n	0.04	.85	n	n	4.39	.04*	n	n
Text reading [b]		78	32.43	.00**	y	n	12.93	.00**	y	n	0.53	.47	n	n
Word spelling [c]		153	12.92	.00**	n	n	0.03	.87	n	n	5.54	.02*	n	n
Go-nogo test mean reaction time per trial														
Word reading[a]	1	121	8.20	.01**	y	n	0.50	.92	n	n	2.48	.12	n	n
	2	120	10.84	.00**	y	n	1.65	.20	n	n	0.99	.32	n	n
	3	111	0.38	.54	n	n	0.36	.55	n	n	0.72	.40	n	n
	4	148	5.35	.02*	n	n	0.15	.70	n	n	0.43	.51	n	n
Text reading [b]		78	24.58	.00**	y	n	6.95	.01*	y	n	0.06	.80	n	n
Word spelling [c]		152	1.20	.28	n	n	0.05	.83	n	n	1.25	.27	n	n
Flanker test letters number of correct responses														
Word reading[a]	1	86	1.34	.25	y	n	2.28	.14	n	n	0.04	.85	n	n
	2	85	6.23	.02*	y	n	3.36	.07	n	n	1.77	.19	y	n
	3	78	0.46	.50	n	n	0.78	.38	n	n	0.42	.52	n	n
	4	104	1.14	.29	n	n	0.35	.55	n	n	0.36	.55	n	n
Text reading [b]		62	2.21	.14	y	n	5.67	.02*	y	n	0.13	.72	n	n

Word spelling [c]		108	6.00	.02*	n	n	0.84	.36	n	n	2.04	.15	n	y
Flanker test letters mean reaction time per trial														
Word reading[a]	1	86	9.69	.00**	y	n	0.42	.52	n	n	6.93	.01	n	y
	2	85	8.97	.00**	y	n	0.68	.41	n	n	0.03	.87	n	n
	3	78	1.25	.27	n	n	0.70	.41	n	n	1.13	.29	n	n
	4	104	6.75	.01*	n	n	0.01	.91	n	n	2.30	.13	n	n
Text reading [b]		62	10.23	.00**	y	n	5.74	.02*	y	n	4.92	.03*	n	n
Word spelling [c]		108	9.47	.00**	n	n	0.01	.94	n	n	2.40	.13	n	n
Flanker test arrows number of correct responses														
Word reading[a]	1	124	7.02	.01**	y	n	0.57	.45	n	n	0.90	.35	n	n
	2	123	7.08	.01**	y	n	0.34	.56	n	n	0.00	.95	n	n
	3	114	0.20	.66	n	n	0.71	.40	n	n	1.62	.21	n	n
	4	151	4.38	.04*	y	n	0.75	.39	n	n	0.11	.74	n	n
Text reading [b]		77	14.08	.00**	y	n	11.59	.00**	y	n	1.30	.26	n	n
Word spelling [c]		155	1.33	.25	n	y	0.04	.84	n	n	0.71	.40	n	n
Flanker test arrows mean reaction time per trial														
Word reading[a]	1	124	26.26	.00**	y	y	1.12	.29	n	n	4.51	.04*	n	n
	2	123	18.16	.00**	y	n	0.47	.49	n	n	0.03	.86	n	n
	3	114	7.27	.01**	n	n	0.87	.35	n	n	0.15	.70	n	n
	4	151	16.47	.00**	y	n	1.16	.28	n	n	1.70	.20	n	n
Text reading [b]		77	21.94	.00**	y	n	12.92	.00**	y	n	0.47	.49	n	n
Word spelling [c]		155	10.74	.00**	n	n	0.00	.98	n	n	6.73	.01*	n	n
Digit recall														
Word reading[a]	1	131	24.26	.00**	y	n	0.05	.83	n	n	1.64	.20	n	n
	2	130	19.16	.00**	y	n	0.04	.85	n	n	0.35	.55	n	n
	3	120	3.30	.07	n	n	0.12	.72	n	n	1.11	.29	n	n
	4	157	13.05	.00**	y	n	0.00	.97	n	n	4.48	.04*	n	n
Text reading [b]		81	17.83	.00**	y	n	15.14	.00**	y	n	2.33	.13	n	n
Word spelling [c]		161	7.00	.01**	n	n	0.58	.45	n	n	2.43	.12	n	n
Backward recall														
Word reading[a]	1	131	21.73	.00**	y	n	1.29	.26	n	n	3.46	.07	n	n

	2	130	20.68	.00**	y	n	0.66	.42	n	n	0.02	.90	n	n
	3	120	6.76	.01*	n	n	0.17	.68	n	n	1.02	.32	n	n
	4	156	15.60	.00**	y	n	0.02	.89	n	n	5.64	.02*	n	n
Text reading [b]		81	25.72	.00**	y	n	16.02	.00**	y	n	0.89	.35	n	n
Word spelling [c]		160	9.18	.00**	n	n	0.37	.55	n	n	3.29	.07	n	n
Blockrecall														
Word reading[a]	1	131	19.42	.00**	y	n	0.40	.53	n	n	3.66	.06	n	n
	2	130	17.29	.00**	y	n	0.06	.81	n	n	0.36	.55	n	n
	3	120	3.33	.07	n	n	0.20	.65	n	n	1.10	.30	n	n
	4	156	15.21	.00**	y	n	0.53	.47	n	y	3.92	.05*	n	n
Text reading [b]		81	24.32	.00**	y	n	21.20	.00**	y	n	2.24	.14	n	n
Word spelling [c]		160	7.19	.01**	n	n	0.23	.64	n	n	2.41	.12	n	n
Naming speed letters														
Word reading[a]	1	133	23.16	.00**	y	n	1.20	.28	n	n	2.84	.09	n	n
	2	132	23.79	.00**	y	n	0.19	.66	n	n	0.03	.86	n	n
	3	122	7.38	.01**	n	n	0.24	.62	n	n	0.63	.43	n	n
	4	164	16.81	.00**	y	n	0.05	.82	n	n	4.12	.04*	n	n
Text reading [b]		83	27.54	.00**	y	n	15.80	.00**	y	n	0.89	.35	n	n
Word spelling [c]		168	11.68	.00**	n	n	0.31	.58	n	n	3.30	.07	n	n
Naming speed digits														
Word reading[a]	1	134	14.17	.00**	y	n	4.66	.03*	n	y	0.73	.39	n	n
	2	133	22.00	.00**	y	n	0.09	.77	n	n	0.38	.54	n	n
	3	123	5.60	.02*	n	n	0.40	.53	n	n	0.20	.66	n	n
	4	165	9.81	.00**	y	n	0.28	.60	n	n	2.79	.10	n	n
Text reading [b]		83	25.34	.00**	y	n	12.15	.00**	y	n	1.60	.21	n	n
Word spelling [c]		169	10.90	.00**	n	n	0.20	.66	n	n	13.64	.00**	n	y
Naming speed pictures														
Word reading[a]	1	134	13.69	.00**	y	n	5.44	.02*	n	y	1.10	.30	n	n
	2	133	16.04	.00**	y	n	0.10	.75	n	n	0.01	.91	n	n
	3	123	6.52	.01*	n	n	0.13	.72	n	n	0.01	.93	n	n
	4	165	14.62	.00**	y	n	0.29	.59	n	n	4.28	.04*	n	n
Text reading [b]		83	20.01	.00**	y	n	3.37	.07	y	n	0.26	.61	n	n

Word spelling [c]		169	8.77	.00**	n	n	0.21	.64	n	n	3.11	.08	n	n
Coding														
Word reading[a]	1	119	25.35	.00**	y	n	0.16	.69	n	y	4.06	.05*	n	n
	2	118	18.82	.00**	y	n	0.47	.50	n	n	0.00	.98	n	n
	3	110	5.17	.03*	n	n	0.05	.83	n	n	0.74	.39	n	n
	4	133	10.93	.00**	n	n	0.01	.93	n	n	3.45	.07	n	n
Text reading [b]		71	25.41	.00**	y	n	11.60	.00**	y	n	1.28	.26	n	n
Word spelling [c]		137	7.27	.01**	n	n	0.13	.72	n	n	2.00	.16	n	n
Score double task number of animals														
Word reading[a]	1	127	14.92	.00**	y	n	0.58	.45	n	n	3.15	.08	n	n
	2	126	9.63	.00**	y	y	1.44	.23	n	n	0.15	.70	n	n
	3	118	2.99	.09	n	n	0.09	.76	n	n	0.12	.73	n	n
	4	156	12.22	.00**	y	n	0.02	.90	n	n	4.51	.04*	n	n
Text reading [b]		83	13.98	.00**	y	n	15.84	.00**	y	n	7.17	.01**	n	y
Word spelling [c]		160	7.18	.01*	n	n	0.27	.60	n	n	1.78	.18	n	n
Score double task number of sounds														
Word reading[a]	1	127	19.20	.00**	y	n	1.22	.27	n	y	3.41	.07	n	n
	2	126	17.16	.00**	y	n	0.71	.40	n	n	0.04	.84	n	n
	3	118	5.59	.02*	n	n	0.03	.86	n	n	1.97	.16	n	n
	4	156	14.18	.00**	y	n	0.01	.93	n	n	4.82	.03*	n	n
Text reading [b]		83	30.73	.00**	y	y	16.32	.00**	y	n	0.87	.35	n	y
Word spelling [c]		160	11.80	.00**	n	n	0.56	.45	n	n	3.77	.05	n	y
Opposite worlds inhibition score														
Word reading[a]	1	128	10.18	.00**	y	n	3.25	.07	n	n	3.12	.08	n	n
	2	127	16.32	.00**	y	n	0.51	.48	n	n	0.00	.97	n	n
	3	119	4.93	.03*	n	n	0.03	.86	n	n	0.22	.64	n	n
	4	157	8.36	.00**	y	n	0.39	.53	n	n	2.80	.10	n	n
Text reading [b]		83	15.15	.00**	y	n	3.87	.05	y	n	0.51	.48	n	n
Word spelling [c]		161	12.67	.00**	n	n	0.52	.47	n	n	3.54	.06	n	n

Note. * $p < .05$ ** $p < .01$. Yellow marked *p*-values indicate change from significance to nonsignificance compared to Table 5; purple marked *p*-values indicate change from nonsignificance to significance compared to Table 5.
[a] Evaluated in terms of total count of correct words. [b] Evaluated in terms of total time required to read the text. [c] Evaluated in terms of total count of correct words. [c] y indicates a significant interaction effect $p < .05$; n indicates a non-significant interaction effect $p > .05$. [d] FU means Follow-up

The introduction of the number of correct responses on the Berg's card sorting test as a covariate, only resulted in a change from significant effects into nonsignificant effects on word reading 3 during the first period of reading and spelling intervention and word reading 1 during the third period of reading and spelling intervention. We found some significant interaction effects of this EF-measure on word reading 3 and 4 during the first period of reading and spelling intervention and on word reading 1 during the third period of reading and spelling intervention. Mean reaction time per trial on the Berg's card sorting test as a covariate only affected the effect of the first period of reading and spelling intervention on word reading 3 and the effect of the third period of reading and spelling intervention on word reading 1: Both effects turned nonsignificant. There is interaction between the mean reaction time per trial on the Berg's card sorting test and text reading in the first and the second period of reading and spelling intervention.

Introducing the number of omission errors on the Go-nogo test only influenced the effect of the third period of reading and spelling intervention on word reading 1, which turned from a significant to an nonsignificant effect. We found one significant interaction effect of this EF-measure on word reading 2 during the third period of reading and spelling intervention. The number of commission errors of the same test as covariate revealed that only the effect of the third period of reading and spelling intervention on word reading 1 turned from a significant effect into an nonsignificant effect, and that the effect on text reading turned from a nonsignificant effect into a significant effect. We found significant interaction effects of this EF-measure on word reading 1 and 2, and text reading during the first period of reading and spelling intervention and on word reading 1 during the second period. The mean reaction time per trial of this test introduced as covariate showed only three *p*-values turned from significant values into nonsignificant one: The effect of the of the first period of reading and spelling intervention on word reading 3 and word spelling, and the effect of the intervention in the third period on word reading 1. Of this EF-measure, no significant interaction effects were found.

The introduction of the number of correct responses of the Flanker test letters as a covariate, altered the effect of the dyslexia treatment in the first period on word reading 1, 3, and 4 and on text reading progress. It also altered the effect of in the third period on word reading 1: All *p*-values turned from significant into nonsignificant effects. We found only one significant interaction value of this EF-measure on word spelling during the third period of reading and spelling intervention. The mean reaction time per trial of the Flanker test letters as covariate affected the progress on word reading 3 during the first period of reading and spelling intervention and the progress on word reading 1 and text reading during the third period of reading and spelling intervention: The first two effects turned from significance into nonsignificance, the latter one from nonsignificance to significance. Only one significant interaction effect of this EF-measure was found, on the progress on word reading 1 during the third period of reading and spelling intervention. The number of correct responses on the flanker test arrows covariate resulted in three *p*-values changing from significant to nonsignificant: Word reading 3 and word spelling during the first period of reading and spelling intervention and word reading 1 during the third period. Only one significant interaction effect of this EF-measure was found, i.e., on word spelling during the first period of reading and spelling intervention. Mean reaction time per trial on the Flanker test arrows as covariate did not affect reading and spelling progress during reading and spelling intervention. We found only one significant interaction effect of this EF-measure on word reading 1 during the first period of reading and spelling intervention.

Introducing digit recall as a covariate, affected the progress during the first period of reading and spelling intervention on word reading 3 and the progress during the third period of reading and spelling intervention on word reading 1. These progress turned from significant into nonsignificant effects. Also, it affected the progress on wording reading 4 during the third period of reading and spelling intervention, which turned from nonsignificant into significant. We found no interaction effects for this EF-measure. Backward recall affected the progress on word reading 1 and 4 during the third period of reading and spelling intervention. The progress on word reading 1 turned from a significant effect into nonsignificant, the progress on word reading 4 revealed the reverse. No interaction effects were found for this EF-measure. Block recall introduced as covariate affected the progress during the first period of reading and spelling intervention on word reading 3 and the progress during the third period of reading and spelling intervention on word reading 1, which turned from significant into nonsignificant. Also, it affected the progress on wording reading 4 during the third period of reading and spelling intervention, turning from an nonsignificant effect into a significant one. We found one significant interaction effect on Text reading during the second period of reading and spelling intervention for this EF-measure.

The introduction of Naming speed letters affected the progress on word reading 1 and 4 during the third period of reading and spelling intervention: The progress on word reading 1 turned from significant into nonsignificant, and the progress on word reading 4 revealed the reverse. No interaction effects were found for this EF-measure. Entering Naming speed digits as covariate affected the progress on word reading 1 during the second period, which turned from nonsignificant into significant. During the third period of reading and spelling intervention, progress word reading 1 turned from significant into nonsignficant and word spelling showed the reverse. Naming speed pictures as covariate, affected the progress on word reading 1 and text reading during the second period of reading and spelling intervention. It also affected progress on word reading 1 and 4 during the third period. During the second period, progress on word reading 1 turned from nonsignificance into significance and progress on text reading showed the reverse. During the third period, progress on word reading 1 turned from nonsignificance into significance, progress on word reading 4 revealed the opposite. Introducing Coding as a covariate did not affect the progress on reading and spelling measures during reading and spelling intervention. Only one interaction effect of this EF-measure was found on word reading 1.

The introduction of the number of animals on Score double task altered the progress on word reading 3 during the first period of reading and spelling intervention and the progress on word reading 1 during the third period. Both effects turned from significant into nonsignificant. Also, the progress on word reading 4 and text reading during the third period were affected, because both turned from nonsignificant into significant. We found two significant interaction effects for this EF-measure: On the progress of word reading 2 during the first period of reading and spelling intervention and on the progress of text reading during the third period of reading and spelling intervention. The number of sounds of the same test as covariate affected the progress on word reading 1 and 4 during the third period of reading and spelling intervention: The first effect turned from significance into nonsignificance and the second effect revealed the opposite. Four significant interaction effects of this EF-measure were found, namely with progress of text reading during the first period of reading and spelling intervention, the progress of word reading during the second period of reading and spelling intervention, and with text reading and word spelling during the third period of reading and spelling intervention.

Introducing the inhibition score of Opposite worlds as covariate affected the progress on text reading during the second period of reading and spelling intervention and the progress on word reading 1 during the third period of reading and spelling intervention. Both effects turned from significance into nonsignificance. We did not find any significant interaction effects for this EF-measure.

Taken together, EF-skills do not appear to affect reading and spelling progress during reading and spelling intervention in a meaningful way. In only 9.5% of the cases, progress in reading and spelling during intervention was explained by the concerned EF skill. Only 3.7% of the effects of reading and spelling intervention on reading and spelling progress turned out to be related to differences in the concerned EF skills. In only 5.6% of the analyses, a significant interaction effect of the concerned EF skill on progress of reading and spelling during reading and spelling intervention was found. We believe that these proportions are negligible.

4. Discussion

The results of the present study reveal that EF measures are related to reading and spelling proficiency in dyslexic children, but not to reading and spelling progress during reading and spelling intervention. The present results differ from previous studies and will be discussed in more depth. Furthermore, we will discuss the present results in light of the theoretical context of dyslexia and reading and spelling intervention. Finally, we will discuss some practical implications of the present study.

4.1 The role of EF in Reading and Spelling Progress

The present study was unique in studying the role of a broad scope of EF measures in reading and spelling *progress* during dyslexia treatment, apart from the role of EF in reading and spelling proficiency. Results revealed that many - but not all - measures of EF are associated with reading and spelling proficiency in dyslexic children. Furthermore, few EF measures appeared to be associated with progress in reading and spelling skills during reading and spelling intervention. This provides evidence for the assumption that EF is not affecting reading and spelling proficiencies. Instead, it is more likely, that problems in EF and problems in reading and spelling are caused by the same problems in underlying cognitive processes. Apparently, these processes account for development in EF as well as for reading and spelling. Therefore, EF of children cannot predict future progress in reading and spelling. This means that reading and spelling problems cannot be explained by EF problems and that it is unlikely that training of EF during reading and spelling intervention will be effective to enhance reading and spelling.

The results of the present study are not entirely in line with the two previous studies on the role of EF in reading and spelling skills on a longitudinal basis (Altemeier et al., 2008; Chung & McBride-Chang, 2011). Chung and McBride-Chang (2011) established that EF-skills predicted Chinese-speaking children's early reading performance to some extent, depending on the regression model that was used. However, their results revealed no significant indications for EF-skills predicting reading development. Altemeier et al. (2008) showed that individual differences in EF predicted literacy outcomes for children with and without dyslexia. There are two reasons for the differences in outcomes from the present study and those of Chung and McBride-Chang (2011) and Altemeier et al. (2008): 1) the EF-measures and 2) the methods of data-analysis that were used.

With respect to the EF measures, the previous studies used a limited number of tasks. Chung and McBride-Chang (2011) only used an inhibitory control task and a working memory task). Altemeier et al., (2008) used only three tasks, namely, a task measuring inhibition, inhibition-switching, and rapid automatic-switching. However, in the present study fifteen different EF-tasks resulting in a total of 21 variables were used in order to evaluate EF. The advantage of such a large test battery is that results are no longer depending on effects that appear coincidentally as a result of the choice of the EF measure. This is important, because there is still a lack of agreement on the theoretical construct of EF and on the choice of tasks that should be used when measuring EF.

With respect to methods of data-analysis, Altemeier et al. (2008) and Chung and McBride-Chang (2011) used multiple regression analysis for analyzing their longitudinal data. However, these studies used literacy outcome as dependent variables, rather than literacy development. Therefore, it was still uncertain whether problems with EF are merely correlated with problems in literacy or whether they can be seen as a cause of literacy problems. In the present study, we accounted for this statistical problem by using four measurements of literacy to model literacy progress and by using repeated measures ANOVA without and with EF-variables as covariates. The advantage of this method is, that the influence of EF on progress in literacy skills over a longer period of time can be investigated. By doing so, it can be concluded that it is more likely that EF does not influence literacy development, even though EF measures and literacy outcomes are correlated.

4.2 Dyslexia in the context of cognitive skills

The results of the present study on the role of EF in reading and spelling progress showed some similarities with previous results on the role of intelligence in reading and spelling development. These previous results showed that it is important to consider the way in which intelligence is evaluated, because intelligence tasks also appeal to expressive language, working memory, and vocabulary (Siegel, 1988). Assessment of an intelligence task puts individuals with reading and spelling disabilities at a disadvantage, because their actual capacities are underestimated. This was illustrated by studies on the genetics of literacy skills and intelligence that showed that the strengths of relationships depend on the specific skills that were tested. For example, a twin study on reading disability revealed that the correlation between intelligence and reading was only weak to moderate, with reading comprehension showing a somewhat stronger relationship (r = .47) than word decoding (r = .38). Moreover, spelling showed a negligible relationship with intelligence (r = .13; Brooks, Fulker, & DeFries, 1990). Thus, the role of intelligence in reading and spelling development appears to resemble that of executive functions in that results depend on the type of measurement instruments that are used and that relationships appear weak to moderate.

Another study that illustrates resemblance of the role of IQ and EF in reading and spelling development showed moderate correlations between intelligence and reading level too. However, the correlation between word decoding and verbal intelligence (r = .46) was somewhat stronger than the correlation between word decoding and performance intelligence (r = .27; Cardon, Dilalla, Plomin, DeFries, & Fulker, 1990). Both studies found that the correlations between intelligence and literacy skills were mostly genetic in origin, which led to the assumption that performance in intelligence tests and literacy measures may be related by the same underlying factors that are genetically determined (Brooks et al., 1990; Cardon et al., 1990).

The relationship between literacy development and intelligence appears to be different for different types of readers. In typical readers, reciprocal influences between reading and intelligence were seen in reading trajectories. In poor readers, however, intelligence influenced weakly but positively the reading scores, whereas reading abilities did not influence intelligence (Ferrer, Shaywitz, Holahan, Marchione, & Shaywitz, 2010).

Thus, the present results on the relation between EF and literacy development show three important domains that resemble outcomes of previous studies on the relationship between intelligence and literacy development. First, it appears that the magnitude of the results depends on the measurement that is used. Second, weak to moderate correlations with literacy skills suggest additional relationships with other underlying cognitive variables. Third, in research on the relation with literacy development of both EF and intelligence, the assumption can be made that underlying processes produce bidirectional relationships between literacy skills on the one hand and other cognitive skills on the other.

The results of the present study fit into an idea of dyslexia (and other learning disorders) as the result of a complex system of developing skills and functions, rather than the idea of dyslexia being caused merely by a phonological deficit (Vellutino et al., 2004). In many cases, reading and spelling failure is accompanied by other cognitive deficits, like sensory deficiencies (Ramus, 2003), motor/coordination deficiencies (Ramus, 2003; Rochelle & Talcott, 2006), working memory (Gathercole et al., 2006), processing speed (Christopher et al., 2012), and attention (Franscheschini et al., 2012).

As illustrated in the present study, reading and spelling proficiency in dyslexic children is associated with EF measures too. Several authors regard a phonological deficit as the primary cause of dyslexia, which is in many cases accompanied by other cognitive deficits (Pennington, 2006; Snowling, 2012; Snowling, 2013). These studies are predominantly based on research regarding co-morbidity with other disorders, like language impairments, symptoms of inattention, attention deficit hyperactivity disorder (ADHD), and motor coordination. Also, some authors propose the idea of different subtypes of dyslexia, with different patterns of affected cognitive functions (Heim & Grande, 2012). Thus, the dominating view of dyslexia is that it is a reading and spelling disorder mainly based on phonological deficiencies, and in many cases is accompanied by other cognitive deficits.

However, the fact that dyslexia is often accompanied by other cognitive deficits in our view, questions the causal role of a phonological deficit. The idea of dyslexia being caused merely by a phonological deficit could be misleading. There’s no doubt that reading development is associated with development of phonological processing skills, as illustrated in the introduction (Vellutino et al., 2004), but correlation does not imply causation. In the correlational studies concerning the relationship between reading skills and phonological processing on which the Vellutino et al. (2004) summary is based, it is often assumed that the correlations that are found, represent phonological processing skills influencing reading development.

In a similar line, it is assumed that weak phonological processing skills hamper reading development. But what about the reverse effect? Morais, Cary, Alegria, and Bertelson (1979) have shown that reading development strongly affects phoneme awareness. The assumption that the relationship between phonological skills and reading is bidirectional seems more likely. Evidence is also provided in a study of early literacy development in children from Kindergarten through Grade 2 (Wagner, Torgesen, & Rashotte, 1994).

A fourth possibility is that the correlations are established by a third factor influencing both phonological processing skills and reading development. The same issues about how to interpret correlations between phonological processing skills and reading development have been proposed before (Stanovich, 1986). It has also been suggested that the nature of the correlation between phonological processing skills and reading development may change with age because cognitive development is a natural maturation process (Stanovich, 1986). In short, using correlational methods it cannot be claimed that some variables should be identified as causes and others as effects. Thus, it is more likely that phonological processing skills are related to reading development to a certain extent, but that they cannot be viewed as causal factors in reading development.

For a final view on the relationships among cognitive skills, we would like to introduce a perspective from systems thinking. Language, reading, and spelling have evolved late in human evolutionary history and require a high level of brain development. These skills are most likely the result of massive interactions of many different brain structures and functions as well as bodily experiences, arriving at any phenotype of reading and spelling abilities (Price, 2000). This line of thinking has been proposed and studied in recent work on children with and without dyslexia (i.e., Holden, Greijn, van Rooy, Wijnants, & Bosman, 2013; Wijnants, Hasselman, Cox, Bosman, & Van Orden (2012).

Holden et al. (2013) presented children with and without dyslexia a reading task (speeded word naming) and three non-reading tasks (simple addition, flanker task, and color naming). In each task, participants were presented with 560 stimuli. The response time on each task generally revealed slower, more variable and skewed (longer right tail) distributions in children with dyslexia than in children without. The difference between the two groups was largest in the speeded word naming and smallest in the flanker task. The fact that children with dyslexia showed distinctly different distributions than children without dyslexia on non-reading tasks as well justified their conclusion that reading is the result of more general perceptual, neurophysiological, biological processes interacting with literacy-learning related factors, such as, poverty, and quality of instruction, rather than the absence or impairment of some dedicated reading process. Despite the fact that children with dyslexia could be distinguished from children without, based on the shape of the response distributions, there was no reason to interpret this as a categorical or qualitative distinction among the two reader groups. The distributions were clearly overlapping and unimodal, showing that the overall shape of the distributions could be rescaled in a self-similar fashion.

We propose a similar explanation for the fact that EF tasks are weakly associated with reading and spelling level. Reading fluency appears to be associated with dynamical instabilities in several cognitive factors that mutually affect one another other. Holden et al. (2013) argued that these instabilities may be due to so-called non-optimal processes in the brain caused by a less effective learning history. The findings concerning EF in the present study may well be understood in light of this perspective.

4.3 Practical Implications

The results of this study reveal that measures of EF show little associations with progress on reading and spelling during reading and spelling intervention. Therefore, reading and spelling intervention should merely focus on training of reading and spelling skills and it is not likely that strengthening EF skills enhances the effect of training reading and spelling.

We endorse to provide children with reading and spelling difficulties an evidence-based reading and spelling remediation program, as outlined in an analysis of The National Reading Panel (2014) in the United States. According to this analysis, an effective reading remediation program should incorporate instruction in phonemic awareness, systematic phonics instruction, improving reading fluency, and enhancement of reading comprehension.

Now it seems that the role of EF in the course of reading remediation is absent, information regarding these skills are unlikely to provide the reading specialist with relevant information. Of course, it helps the intervention process if a child can sit still and concentrate, but it is not an indication of the potential of becoming a better reader. It is up to the reading specialist, the teacher, and perhaps the parents to enhance the circumstances in which a child has to learn, requiring social skills rather than knowledge about reading intervention. Information from the parents and the teacher will provide the reading specialist with knowledge about the way in which to approach the child so that reading instruction will fall on fertile soil.

5. References

Altemeier, L.E., Abbot, R.D. & Berninger, V.W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *Journal of Clinical and Experimental Neuropsychology, 30,* 588-606. doi: 10.1080/13803390701562818

Banich, M.T. (2009). Executive function. The search for an integrated account. *Current Directions in Psychological Science, 18*, 89-94. doi: 10.1111/j.1467-8721.2009.01615.x

Barkley, R.A. (1996). Critical issues in research on attention. In G.R. Lyon, & N.A. Krasnegor, (Ed.), *Attention, Memory, and Executive Function* (pp. 235-278). Baltimore, MD: Brookes.

Berninger, V.W., Abbott, R.D., Thomson, J., Wagner, R., Swanson, H., Wijsman, E.M. & Raskind, W. (2009). Modeling phonological core deficits within a working memory architecture in children and adults with developmental dyslexia. *Scientific Studies of Reading, 10:2,* 165-198. doi: 10.1207/s1532799xssr1002_3

Best, J.R., Miller, P.H. & Jones, L.L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29,* 180-200. doi: 10.1016/j.dr.2009.05.002

Bezdjian, S., Baker, L.A., Lozano, D.I. & Raine, A. (2009). Assessing inattention and impulsivity in children during the Go/NoGo task. *British Journal of Developmental Psychology, 27*, 365-383. doi: 10.1348/026151008x314919

Blomert, L. & Vaessen, A. (2009). *3DM Differentiaal Diagnostiek van Dyslexie. Cognitieve Analyse van Lezen en Spellen. Gebruikershandleiding*. Amsterdam, the Netherlands: Boom.

Booth, J.N., Boyle, J.M.E. & Kelly, S.W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *British Journal of Psychology, 28,* 133-176. doi: 10.1348/026151009x485432

Borkowski, J.G. & Burke, J.E. (1996). Theories, models, and measurements of executive functioning. In G.R. Lyon, & N.A. Krasnegor, (Ed.), *Attention, Memory, and Executive Function* (pp. 235-278). Baltimore, MD: Brookes.

Braams, T. (2008). Dutch dyslexia policies: The ‘dyslexia masterplan’. In Singleton, C. (Ed.), *The Dyslexia Handbook 2008/9.* Shaftesbury, United Kingdom: Blackmore.

Brooks, A., Fulker, D.W. & DeFries, J.C. (1990). Reading performance and general cognitive-ability – A multivariate genetic-analysis of twin data. *Personality and individual differences, 2,* 141-146. doi: 10.1016/0191-8869(90)90006-d

Brosnan, M., Demetre, J., Hamill, S., Robson, K., Sheperd, H. & Cody, G. (2002). Executive functioning in adults and children with developmental dyslexia. *Neuropsychologia, 40,* 2144-2155. doi: 10.1016/s0028-3932(02)00046-5

Brus, B. Th. & Voeten, M.J.M. (1972). *EenMinuutTest. [One Minute Test]* Nijmegen, the Netherlands: Berkhout.

Cardon, L.R., Dilalla, L.F., Plomin, R., DeFries, J.C. & Fulker, D.W. (1990). Genetic correlations between reading performance and IQ in the Colorado adoption project. *Intelligence, 14,* 245-257. doi: 10.1016/0160-2896(90)90008-h

Christopher, M.E., Miyake, A., DeFries, J.C., Wadsworth, S.J., Willcutt, E., Keenan, J.M., Pennington, B. & Olson, R.K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology, 141,* 470-488. doi: 10.1037/a0027375

Chung, K.K.H. & McBride-Chang, C. (2011). Executive functioning skills uniquely predict Chinese word reading. *Journal of Educational Psychology, 103,* 909-921. doi: 10.1037/a0024744

Damasio, H., Grabowski, T., Frank, R., Gallaburda, A.M. & Damasio, A.R. (1994). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science, 264,* 1102-1105. doi: 10.1126/science.8178168

Davelaar, E.J. & Stevens, J. (2009). Sequential dependencies in the Eriksen flanker task: A direct comparison of two competing accounts. *Psychonomic Bulletin and Review, 16,* 121-126. doi: 10.3758/pbr.16.1.121

De Weerdt, F., Desoete, A. & Roeyers, H. (2013). Behavioral inhibition in children with learning disabilities. *Research in Developmental Disabilities, 34,* 1998-2007. doi: 10.1016/j.ridd.2013.02.020

Duncan, J., Burgess, P. & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia, 33*, 261-268. doi: 10.1016/0028-3932(94)00124-8

Duncan, J., Emslie, H., Williams, P., Johnson, R. & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30,* 257-303. doi: 10.1006/cogp.1996.0008

Feldman, R.P & Goodrich, J.T. (2001). Psychosurgery: A historical overview. *Neurosurgery, 48,* 647-659. doi: 10.1097/00006123-200103000-00041

Ferrer, E., Shaywitz, B.A., Holahan, J.M., Marchione, K. & Shaywitz, S.E. (2010). Uncoupling of reading and IQ over time. *Psychological Science, 21,* 93-101. doi: 10.1177/0956797609354084

Foy, J.G., & Man, V.A. (2013). Executive function and early reading skills. *Reading and Writing, 26,* 453-472. doi: 10.1007/s11145-012-9376-5

Franscheschini, S., Gori, S., Ruffino, M., Pedrolli, K. & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology, 22,* 1-6. doi: 10.1016/j.cub.2012.03.013

Gathercole, S.E., Alloway, T.P., Willis, C. & Adams, A.M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*, 265-281. doi: 10.1016/j.jecp.2005.08.003

Geelhoed, J. & Reitsma, P. (2004). *PI-dictee.[PI-dictation]* Amsterdam, the Netherlands: Harcourt.

Gerschberg, F.B. & Shimamura, A.P. (1995). Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia, 13*, 1305-1333. doi: 10.1016/0028-3932(95)00103-A

Goldstein, L.H., Bernhard, S., Fenwick, P.B., Burgess, P.W. & McNeill, J. (1993). Unilateral frontal lobectomy can produce strategy application disorder. *Journal of Neurology, Neurosurgery and Psychiatry, 56,* 274-276. doi: 10.1136/jnnp.56.3.274

Halperin, J.M., Wolf, L., Greenblatt, R. & Young, G. (1991). Subtype analysis of commission errors on the continuous performance test in children. *Developmental Neuropsychology, 7,* 207-217. doi: 10.1080/87565649109540488

Heim, S. & Grande, M. (2012). Fingerprints of developmental dyslexia. *Trends in Neuroscience and Education, 1,* 10-14. doi: 10.1016/j.tine.2012.09.001

Heim, S., Tschierse, J., Amunts, K., Wilms, M., Vossel, S., Willmes, K., Grabowska, A. & Huber, W. (2008). Cognitive subtypes of dyslexia. *Acta Neurobiologiae Experimentalis, 68*, 73-82.

Helland, T. & Asbjornsen, A. (2000). Executive functions in dyslexia. *Child Neuropsychology, 6*, 37-48. doi: 10.1076/0929-7049(200003)

Holden, J.G, Greijn, L., van Rooij, M.J.W., Wijnants, M.L., & Bosman, A.M.T. (2013). *Cognitive dynamics of dyslexia*. Manuscript submitted for publication.

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development, 20,* 251-271. doi: 10.1002/icd.736

Jeffries, S. & Everatt, J. (2004). Working memory: its role in dyslexia and other specific learning difficulties. *Dyslexia, 10,* 196-214. doi: 10.1002/dys.278

Jongen, I. & Krom, R. (2009). *Drie Minuten Toets (DMT) [Three Minute Test]*. Arnhem, the Netherlands: CITO B.V.

Jurado, M.B. & Rossellini, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychological Review, 17*, 213-233. doi: 10.1007/s11065-007-9040-z

Kegel, C.A.T. & Bus, A.G. (2013). Links between DRD4, executive attention, and alphabetic skills in a nonclinical sample. *Journal of Child Psychology and Psychiatry, 54,* 305-312. doi: 10.1111/j.1469-7610.2012.02604.x

Levine, D.S. & Prueitt, P.S. (1989). Modeling some effects of frontal lobe damage – novelty and perseveration. *Neural Networks, 2*, 103-116. doi: 10.1016/0893-6080(89)90027-0

Lyon, G.R. (1996). The need for conceptual and theoretical clarity in the study of attention, memory, and executive function. In G.R. Lyon, & N.A. Krasnegor, N.A. (Ed.), *Attention, Memory, and Executive function* (pp. 235-278). Baltimore, MD: Brookes.

Manly, T., Robertson, I.H., Anderson, V. & Nimmo-Smith, J. (1999/2004). *TeaCh. Test of Everyday Attention for Children.* London, United Kingdom: Harcourt Assessment.

Marzocchi, G.M., Oosterlaan, J., Zuddas, A., Cavolina, P., Geurts, H. Redigolo, D., …, Sergeant, J.A. (2008). Contrasting deficits on executive functions between ADHD and reading disabled children. *Journal of Child Psychology and Psychiatry, 49,* 543-552. doi: 10.1111/j.1469-7610.2007.01859.x

Menghini, D., Finzi, A., Benassi, M., Bolzani, R., Facoetti, A., Giovagnoli, S., …, Vicari, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: A comparative study. *Neuropsychologia*, *48,* 863-872. doi: 10.1016/j.neuropsychologia.2009.11.003

Miyake, A. & Friedman, N.P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21:1*, 8-14. doi: 10.1177/0963721411429458

Miyake, A. Friedman, N.P. Emerson, M.J., Witzki, A.H. Howerter, A. & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*, 49-100. doi: 10.1006/cogp.1999.0734

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323-331. doi: 10.1016/0010-0277(79)90020-9

Mueller, S.T. (2010). PEBL: Psychology Experiment Building Language. Retrieved from http://pebl.sourceforge.net/

Nauta, W.J.H. (1971). The problem of the frontal lobe: A reinterpretation. *Journal of Psychological Research, 8,* 167-187. doi: 10.1016/0022-3956(71)90017-3

National Reading Panel (2014). *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Literature on Reading and Its Implications for Reading Instruction.* Retrieved from http://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Nationaal Referentiepunt Dyslexie (NRD). (2013). *Protocol Dyslexie Diagnostiek & Behandeling 2.0* Retrieved from http://www.nrd.nu/images/downloads/verantwoording_PDDB2.0.pdf

Niendam, T.A., Laird, A.R., Ray, K.L., Dean, Y.M., Glahn, D.C. & Carter, C.S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective and Behavioral Neuroscience, 12,* 241-268. doi: 10.3758/s13415-011-0083-5

Obonsawin, M.C., Crawford, J.R., Page, J., Chalmers, P. & Cochrane, R. (2002). Performance on tests of frontal lobe function reflect general intellectual ability. *Neuropsychologia,40,* 970-977. doi: 10.1016/s0028-3932(01)00171-3

Packwood, S., Hodgetts, H.M. & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *Journal of Clinical and Experimental Neuropsychology, 33*, 456-470. doi: 10.1080/13803395.2010.533157

Pennington, B.F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101*, 385-413. doi:10.1111/j.1469-7610.1996.tb01380.x

Pennington, B.F., Benetto, L., McAleer, O. & Roberts, R.J. (1996). Executive functions and working memory: Theoretical and measurement issues. In G.R. Lyon, & N.A. Krasnegor (Ed.), *Attention, Memory, and Executive Function* (pp. 235-278). Baltimore, MD: Brookes.

Pennington, B.F. & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37,* 51-87. doi: 10.1016/j.cognition.2006.04.008

Pickering, S. & Gathercole, S. (2001). *Working Memory Test Battery for Children (WMTB-C).* London, United Kingdom: The Psychological Corporation Limited.

Poljac, E., Simon, S., Ringlever, L., Kalcik, D., Groen, W.B., Buitelaar, J.K. & Bekkering, H. (2010). Impaired task switching performance in children with dyslexia but not in children with autism. *The Quarterly Journal of Experimental Psychology, 63,* 401-416. doi: 10.1080/17470210902990803

Price, C.J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy, 197,* 335-359. doi: 10.1046/j.1469-7580.2000.19730335.x

Protopapas, A. Archonti, A. & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology, 54*, 251-282. doi: 10.1016/j.cogpsych.2006.07.003

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P.& Baayen, H. (2013). The myth of cognitive decline: Non-Linear dynamics of lifelong learning. *Topics in Cognitive Science, 6*, 5-42. doi: 10.1111/tops.12078

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current opinion in neurobiology, 13*, 212-218. doi: 10.1016/s0959-4388(03)00035-7

Ramus, F. & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology, 61*, 129-141. doi: 10.1080/17470210701508822

Reiter, A., Tucha, O. & Lange, K.W. (2004). Executive functions in children with dyslexia. *Dyslexia, 11,* 116-131. doi: 10.1002/dys.289

Rochelle, K.S.H. & Talcott, J.B. (2006) Impaired balance in developmental dyslexia? A meta-analysis of the contending evidence. *Journal of Child Psychology and Psychiatry, 47*, 1159-1166. doi: 10.1111/j.1469-7610.2006.01641.x

Royall, D.R., Lauterbach, A.C., Cummings, J.L., Reeve, A., Rummans, T.A., Kaufer, D.I., …, Coffey, C.E. (2002). Executive control function: A review of its promise and challenges for clinical research. *Journal of Neuropsychiatry and Clinical Neuroscience, 14,* 337-405. doi: 10.1176/appi.neuropsych.14.4.377

Salthouse & Davis (2006). Organization of cognitive abilities and neuropsychological variables across the lifespan. *Developmental Review, 26,* 31-54. doi: 10.1016/j.dr.2005.09.001

Shanahan, M.A., Pennington, B.F., Yerys, B.E., Scott, A., Boada, R., Willcutt, …, DeFries, J.C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*, 585-602. doi: 10.1007/s10802-006-9037-8

Siegel, L.S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. *Canadian Journal of Psychology, 42*, 201-215. doi: 10.1037/h0084184

Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford, United Kingdom: University Press.

Smith, E.E. & Jonides, J.J. (1999). Storage and executive processes in the frontal lobes. *Science, 283,* 1657-161. doi: 10.1126/science.283.5408.1657

Snowling, M.J. (2012). Editorial: Virtual issue. Changing concepts of dyslexia: nature, treatment and comorbidity. *Journal of Child Psychology and Psychiatry*, *53*, 1-3. doi: 10.1111/j.1469-7610.2009.02197.x

Snowling, M.J. (2013). Early identification and interventions for dyslexia: A contemporary view. *Journal of Research in Special Educational Needs, 13*, 7-14. doi: 10.1111/j.1471-3802.2012.01262.x

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407. doi: 10.1598/rrq.21.4.1

Steinberg, L. (2004). Risk taking in adolescence. What changes and why? *Annals New York academy of sciences, 1021*, 51-58. doi: 10.1196/annals.1308.005

Sternberg, R.J. (1985). *Beyond IQ. A Triarchic Theory of Human Intelligence.* New York, NY: Cambridge University.

Stuss, D.T., Gow, C.A. & Hetherington, C.R. (1992). "No longer Gage". Frontal lobe dysfunction and emotional changes. *Journal of Consulting and Clinical Psychology, 60,* 349-359. doi: 10.1037/0022-006x.60.3.349

Treitz, F.H., Heyder, K. & Daum, I. (2007). Differential course of executive control changes during normal aging. *Aging, Neuropsychology, and Cognition, 14*, 370-393. doi: 10.1080/13825580600678442

Vellutino, F. R., Fletcher, J.M., Snowling, M.J. & Scanlon, D.M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*, 2-40. doi: 10.1046/j.0021-9630.2003.00305.x

Visser, J., Van Laarhoven, A. & Ter Beek, A. (1996). *AVI Toetspakket. [AVI Test Package]* ’s-Hertogenbosch, the Netherlands: KPC.

Wagner, R.K., Torgesen, J.K. & Rashotte, C.A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*, 73-89. doi: 10.1037/0012-1649.30.1.73

Wechsler, D. (1991). Wechsler Intelligence Scale for Children (3rd ed.), San Antonio, TX: Psychological Corporation.

Wechsler, D. (2004). Wechsler Intelligence Scale for Children-Fourth UK Edition (WISC-IV UK). London, United Kingdom: Psychological Corporation.

Wechsler, D. (2005). *WISC-III NL*. London, United Kingdom: Harcourt Assessment.

Wechsler, D. & Naglieri, J.A. (2008). *Wechsler Nonverbal Scale of Ability. Nederlandstalige Bewerking (WNV NL).* Amsterdam, the Netherlands: Pearson Assessment and Information.

Wijnants, M.L., Hasselman, F., Cox, R.F.A., Bosman, A.M.T. & Van Orden, G. (2011). An interaction-dominant perspective on reading fluency and dyslexia. *Annals of Dyslexia, 62*, 100-119. doi: 10.1007/s11881-012-0067-3