Published in final edited form as:

Learn Disabil Q. 2019 February ; 42(1): 17–31. doi:10.1177/0731948717749935.

Executive Functions and Response to Intervention: Identification of Students Struggling with Reading Comprehension

Jeremy Miciak¹, Paul T. Cirino¹, Yusra Ahmed¹, Erin Reid¹, and Sharon Vaughn²

¹University of Houston

²University of Texas at Austin

Abstract

Inadequate responders demonstrate significant risk for learning disabilities. Previous investigations of the cognitive profiles of inadequate and adequate responders have not included measures of executive functions (EF), which have well-documented associations to reading comprehension. We evaluated EF performance on a common factor comprised of shared variance across tasks as well as five separable EF factors in the context of an intensive reading intervention for struggling fourth graders. To determine if EF performance at pretest is associated with subsequent responder status, we compared EF performance of three subgroups of students: inadequate and adequate responders and typical students not at-risk for reading disabilities. Results of discriminant function analyses and linear regression models comparing groups were largely null; EF performance at pretest demonstrated only small associations with responder status. These results suggest that the assessment of EF may have limited value in predicting which individual students will respond to intensive reading interventions.

Keywords

Executive Function; Reading Comprehension; Response to Intervention; Identification

We define executive functions (EF) as domain general control processes that permit individuals to complete goal-directed activities (Cirino et al., in press), similar to other broad-based definitions (e.g., Miyake et al., 2000; Lezak, Howienson, Bigler, & Tranel, 2012), although it is well-recognized that EF can be characterized in many different manners (Suchy, 2009). As most school-based tasks represent goal-directed activities (e.g. reading or performing mathematics calculations), there is considerable intuitive and theoretical appeal for investigating EF in educational contexts. Performance on measures of EF is strongly associated with academic performance in key academic areas, such as reading, writing, and math (Blair & Razza, 2007; Gerst et al., 2015; Yeniad, Malda, Messman, van Ijzendoorn, & Pieper 2013), and deficits in EF are associated with difficulties in reading (e.g., working memory; Cain, Oakhill, & Bryant, 2004).

The term EF has its roots in neuropsychology. For example, Goldstein (1949) noted that patients who had sustained damage to the frontal lobes of the brain had a deficit in “abstract attitude” with effects on both performance tests as well as everyday behavior, although earlier case studies were available (e.g., Harlow, 1848). Pribram (1973) was one of the first to use the words “executive” in connection with the frontal cortex. Since that time, the conceptualization of EF has become more expansive, in several ways. First, the meaning of EF has become decoupled from a one-to-one relationship with frontal lobe functioning (Stuss, 1992), even as this relationship continues to be explored. Second, the structure of EF (how its components relate to one another) is becoming more clear across age ranges (e.g., Cirino et al., in press; Miyake et al., 2000; Weibe et al., 2011), and there is now greater information about how EF components such as working memory, inhibition, shifting, planning, etc. relate to one another. Further, EF is known to develop and improve throughout childhood (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Best, Miller, & Jones, 2009), including in terms of its structure (e.g., Huizinga, Dolan, & van der Molen, 2006; Lee, Bull, & Ho, 2013), all of which is consistent with neurobiological development, particularly with regard to the frontal lobes (Jurado & Rosselli, 2007). Third, EF has come to be associated with other broad functions, each of which has their own historical traditions across cognitive, developmental, and education literatures. These functions include attention, working memory, metacognition, and self-regulation, and there is recognition of the relationship of EF to such skills (Hoffmann, Schmeichel, & Baddeley, 2012), assessed either with laboratory tasks or rating scales. EF has been associated with achievement outcomes at the level of both intervention and prediction. For example, intervention approaches have focused on either broad, curricular approaches that emphasize self-regulation (Diamond & Lee, 2011) or explicit working memory training (Melby-Lervag & Hulme, 2013), although the effects of neither have been robust. Prediction studies are common, though it is rarer to couple this with explicit achievement intervention. The present paper attempts to both use a conceptualization of EF that reflects its complexity, across the manners in which it is assessed (see below), and also takes place in the context of a reading intervention that has a self-regulatory component.

The well-established associations between academics and EF has spurred interest in whether interventions targeting EF (broadly construed) may result in improved academic performance and improvement in reading (e.g. Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Diamond, Barnett, Thomas, & Munro, 2007). This issue was investigated in a recent meta-analysis (Jacob & Parkinson, 2015) that evaluated both the unconditional relationship of EF to academic skills, as well as causal relations – i.e., whether manipulating EF experimentally results in improved academic performance. Jacob and Parkinson (2015) concluded that there was a clear relationship between academic performance and EF in its various instantiations, but found no compelling evidence that associations between EF and academic performance are *causal*. Of note, Jacob and Parkinson identified only five studies that included a robust set of covariates (e.g. maternal education, socio-economic status, or prior achievement) deemed sufficient to suggest a *unique* role for EF in academic performance; among the 13 correlations between EF and academics in these studies, only one, a correlation between working memory and math achievement (Fitzpatrick & Pagani, 2012), was significant and positive.

Beyond correlational or predictive relationships, and experimental manipulation of EF, there is a *third* potential way for EF to impact reading performance. EF may impact the manner in which students respond to intensive reading interventions. For example, a threshold of skill in EF prior to participating in the intensive intervention may be necessary, if not sufficient, for responding effectively to the reading intervention. Alternatively, students with differing levels of EF may respond differently to reading intervention, with good EF potentially compensating to some extent for skill deficits. These possibilities are especially plausible for older students (i.e., beyond grade 2), for whom the focus of reading tasks is on comprehension rather than or in addition to a word level focus. In this context, EF may be utilized to help scaffold or control essential processes of reading comprehension, such as incorporating background knowledge, summarizing passages, or monitoring engagement with text. To our knowledge, previous reviews of factors associated with intervention response in reading have not included studies explicitly evaluating EF or those focusing on reading comprehension in older students (Stuebing et al., 2015), which makes this an innovative area of inquiry. To address such questions, studies must include measures of EF, academic achievement, and include (and experimentally manipulate) intensive reading intervention. Below, we describe the importance of studying the relations between EF and reading intervention response in the context of intensive interventions, and the EF characteristics of students who do and who do not demonstrate adequate response to intensive interventions. The key question that guides this study is, *Can baseline (pretest) measures of EF help us understand which individual students will respond to intensive reading interventions?*

EF and Intervention Response

The importance of intervention response.

There are many potential causes of poor reading, which confounds generalizations from many cross-sectional and correlational research studies that include cognitive and neuropsychological dimensions. For example, academic difficulties may be the result of limited academic opportunity or ineffective instruction (Fletcher & Vaughn, 2009; Fuchs, Fuchs, & Vaughn, 2008). Students arrive at school with different language and literacy experiences (Bhattacharya, 2010; Molfese & Molfese, 2002). Students engage in heterogeneous language and literacy activities (Connor & Morrison, 2006; Foorman et al., 2006; Piasta, Connor, Fishman, & Morrison, 2009) requiring consideration not only of reading level among students with reading difficulties, but also whether students have received adequate opportunity, in the form of intensive, evidence-based reading interventions (Fletcher et al., 2011; Miciak et al., 2014).

The attributes of inadequate responders: A critical subgroup.

Studies that focus on intervention response permit inferences about inadequate responders, a population of considerable interest in special education practice and research and who by definition demonstrate intractable academic difficulties. In the last decade and a half, there has been movement towards learning disabilities (LD) identification criteria that rely, in part, on inadequate response to intensive interventions as an inclusionary criterion, data which are typically collected as part of a school-wide response to intervention framework (also called

multi-disciplinary systems of support; Gresham et al., 2005; VanDerHeyden & Burns, 2010). This procedural shift mandating that students identified with LD must demonstrate inadequate response to evidence-based interventions has not been without controversy, as some have questioned the validity of LD identification decisions emerging from RTI frameworks (Hale et al., 2010; Kavale, Kauffman, Bachmeier, & LeFever, 2008). However, the mandate for documentation of inadequate response to intervention is not unique to methods that rely on RTI frameworks. Across all potential LD identification criteria permitted by IDEA 2004, the comprehensive assessment process must document inadequate response to evidence-based interventions (Fletcher et al., 2007). This mandate to document intervention response as an inclusionary criterion for all LD identification highlights the critical need to provide empirical evidence for the validity of classifications based on instructional response (Fletcher et al., 2011; Miciak et al., 2014). Students who demonstrate differential response to evidence-based, intensive interventions can be compared on dimensions not utilized for group formation –and EF is a key example of these. If students differ on such external dimensions, the classification accrues validity (Morris & Fletcher, 1988). By focusing on EF, the present study extends validity research on RTI methods by comparing inadequate and adequate responders in this important and previously unstudied domain.

Understanding the attributes of inadequate responders also contributes to research that identifies potential intervention targets. For example, inadequate responders are known to have deficits in several cognitive and linguistic domains, including phonological awareness, vocabulary, and listening comprehension (Fletcher et al., 2011; Stage, Abbott, Jenkins, & Berninger, 2003). Many of these domains are promising intervention targets (Lesaux, Kieffer, Kelley, & Harris, 2014; Torgesen, Wagner, Rashotte, Herron, & Lindamood, 2010). Therefore, if EFs are associated with responder status, then to the extent these domains of EF are malleable, it would serve as another promising intervention target, potentially as an adjuvant to intensive reading interventions.

Predicting who will respond: The importance of accurate prediction.

A second reason to evaluate EFs in the context of reading interventions is that findings from this evaluation may prompt educators to provide more intensive intervention earlier. If measures of EF are demonstrated to be associated with intervention response, then administering these measures early and providing early treatment may provide for more effective outcomes. This is especially relevant because most response to intervention service delivery frameworks include three to four tiers of increasing intensity (Jimerson, Burns, & VanDerHeyden, 2016; Kovalski, VanDerHeyden, & Shapiro, 2013). The idea that cognitive and linguistic measures could be useful predictors of individual response to intensive reading interventions is not new. Previous literature reviews (Al Otaiba & Fuchs, 2002) and meta-analyses (Nelson, Benner, & Gonzalez, 2003) have implicated a number of skills as potential predictors of intervention response, including phonological awareness, rapid naming, problem behavior, and deficits in working memory. However, these reviews provide associations between cognitive attributes and reading problems and do not evaluate the extent to which baseline cognitive performance predicts response status following intensive reading intervention.

Stuebing and colleagues' (2015) meta-analysis evaluating predictors of intervention response identified three models by which predictors of intervention response have been investigated: (1) unconditional growth curve models; (2) unconditional gain models; and (3) change conditioned on initial reading status. The three models address distinct research questions, and confusion about differing analytic frameworks may account for different interpretations of evidence on the relative contributions of student characteristics to intervention response. Model 3 is particularly pertinent in the present context, as it directly addresses a "value-added" component to which the assessment of EF could contribute (Fletcher et al., 2011; Miciak et al., 2014). Academic assessments are inexpensive, easily administered, and directly inform intervention. To justify the use of more comprehensive psychoeducational assessments (e.g. cognitive assessment and assessments of EF), it would be critical to show that they contribute unique and meaningful information about student response—that is that cognitive and EF assessments explain unique variance in posttest outcomes (Fletcher & Miciak, 2017).

In the present study, we identify whether EF contributes unique variance to responder status (inadequate versus adequate), beyond baseline reading variables. Previous studies of the association between academic status and cognitive functioning have demonstrated a clear stepwise progression in which students who experience the most significant academic delays experience the most significant cognitive delays, consistent with a continuum of severity hypothesis (Fletcher et al., 2011; Vellutino et al., 2006).

The Present Study: Context and Purpose

The present study extends a previous investigation of the structure of EF (Cirino et al., in press). In that study, 27 indicator measures were assessed, representing eight potentially separable latent EF factors from across cognitive, neuropsychological, and educational literatures. Multiple models were evaluated; the best fitting model was a bifactor model with a common EF factor, as well as five specific latent factors: (1) working memory as recall span with and manipulation/planning; (2) working memory as ongoing maintenance of information and updating in active memory; (3) generative fluency (e.g. verbal fluency) as efficient and accurate retrieval from semantic memory within a set of parameters; (4) self-regulated learning as the ability to control one's own learning processes, including strategy use and self-efficacy; and (5) metacognition as executive processes that monitor, manipulate, or control/regulate other cognitive processes. This final model serves as the basis for the analyses conducted in the present study. For purposes of the present study, factor scores were generated from the latent model (common EF and five specific EF factors), resulting in scores along a z-score metric. The present study extends Cirino et al. (in press) with previously unreported data on a subsample of struggling readers (Vaughn et al., 2016). These students received an intensive reading intervention and were later empirically classified as either inadequate or adequate responders to the intervention using post-intervention data. These two groups were compared to one another and a group of typical students who did not demonstrate reading difficulties. We then evaluate whether there is unique variance in responder status explained by identified EF factors, after including pretest reading and language performance.

We hypothesized that the results of this study would be consistent with a continuum-of-severity hypothesis. That is, because EF is related to reading, and typical readers outperform adequate responders, who in turn (by definition) outperform inadequate responders, we expect EF to differentiate these groups in a similar step-wise fashion. We expect this differentiation to remain (although diminished in size) once initial reading status is accounted for, with typical readers and adequate responders demonstrating relatively better performance on measures of EF. Given the robustness of the common EF factor, we expect differences on this factor score to be the largest, but it is more difficult to form strong *a priori* hypotheses about a differential pattern of EF to be exhibited by the three groups on the specific EF factors. Our expectations are both bolstered by the robustness of the measurement of the common factor, but also tempered by the evidence with other cognitive functions to date, which have shown little added value (Fletcher et al., 2011; Miciak et al., 2014; Stuebing et al., 2015). Nonetheless, the present study contributes to our understanding of the role that cognitive and neuropsychological processes play in reading outcomes.

Method

Overview

The present study, as noted previously, extends findings from a measurement study on the structure of EFs (Cirino et al., in press) and an intervention study of an intensive reading intervention implemented with struggling readers (Vaughn et al., 2016). Participants for the present study include students who completed the full EF battery from the measurement study and also received intensive reading intervention provided in the intervention study. The study also compared the intervention students to typical students (i.e., students who did not demonstrate reading risk) who were randomly selected as part of the measurement study and were assessed in the fall and spring. The dataset, research questions, and analytic plan for the present study are unique. No previous study investigating the effects of the intensive intervention has analyzed EF as a potential predictor of intervention response.

Participants and Group Formation

School sites.—Participants for the present study were drawn (approximately equally) from two sites: one large urban district (eight elementary schools) and a second site encompassing two close urban school districts (nine elementary schools) in the Southwestern United States. The mean enrollment among the 17 schools was 697 students (range 425 to 1140), and the percent of students who received free or reduced lunch was 81.6% (range 46.1% to 98.4%). All schools were rated as academically acceptable. Students for both the measurement and intervention studies were drawn from the same set of schools and sites.

Selection of intervention participants.—The target population for the intervention was 4th grade students at-risk for reading disabilities, as measured by a standard score of 85 or below on the Gates MacGinitie Reading Test (GMRT; MacGinitie, MacGinitie, Maria, Dryer, & Hughes, 2000). Of the 1,695 fourth graders who were screened across the 17 participating schools, 487 were identified as struggling readers. Demographic information is provided in Table 1. The struggling readers were randomly assigned at the student level

within schools in a 2:1 ratio to either a researcher-provided intervention group or to a business as usual (BAU) comparison group. There were no statistically significant differences between the groups according to participant age, $t(479) = 1.15, p > .05$, free or reduced lunch status, $\chi^2(1) = 0.00, p > .05$, special education status, $\chi^2(1) = 0.34, p > .05$, or race/ethnicity, $\chi^2(1) = 3.20, p > .05$. Of students randomly assigned to the researcher-provided intervention, several attrited (see Vaughn et al., 2016 for a full review), but analysis revealed no differences between attrited and non-attrited students by score on the Gates-MacGinitie Reading Test at pretest $t(482) = 0.05, p > .05$, school type (urban or near urban; $\chi^2(1) = 2.50, p > .05$), or intervention assignment (researcher-delivered intervention or BAU; $\chi^2(1) = 0.76, p > .05$). Additional students did not complete the pre- and post-test academic assessments, resulting in 231 intervention students included in the current study.

Study participants.—The present study includes 305 participants from two nonoverlapping pools: intervention students and typical readers. Intervention students ($n = 231$) were students who were assigned to the intervention condition and completed the pre- and post-test academic assessments, which permitted classification of individual students as responders or inadequate responders. Typical readers ($n = 74$) were students who achieved a standard score greater than 90 at pretest on the primary reading criterion measures (below), completed the EF measures, and completed the post-test academic battery.

Categorization of inadequate responders.—Participants who completed the researcher-provided intervention were classified as inadequate or adequate responders based on posttest performance on two measures of reading comprehension: the Woodcock Johnson Third Edition – Passage Comprehension (WJ-III Passage Comprehension; Woodcock, McGrew, & Mather, 2001) subtest and the Gates MacGinitie Reading Test (GMRT; MacGinitie et al., 2000). Students who received a standard score of 85 or below on the GMRT *or* a score of 86 or below on the WJ-III Passage Comprehension subtest were classified as inadequate responders. Students who scored above 85 on the GMRT *and* above 86 on the WJ-III Passage Comprehension subtest were classified as adequate responders. These cut-points were applied because they represented the threshold for the bottom quartile for both measures based on a local, benchmark sample (described below). The application of these criteria resulted in 167 inadequate responders and 64 adequate responders.

Our decision to employ final-status criteria (as opposed to growth criteria or dual discrepancy criteria; Fuchs & Deshler, 2007) is based on two considerations. First, final status directly addresses a critical educational programming question: *does the student continue to require intensive reading intervention?* Additionally, indices of growth add little predictive power over final status to a determination of adequate response (Schatschneider, Wagner, Crawford, 2008; Tolar, Barth, Fletcher, Francis, & Vaughn, 2013), particularly in samples drawn beyond early elementary grades. This is because indices of growth and final status are highly correlated when selection is based on performance at time point one (i.e. pretest; as in this study). The only manner in which students can achieve differential final status is differential (fan-shaped) growth.

Our decision to utilize multiple, final-status indicators of reading comprehension to determine intervention response matches previous studies of the cognitive characteristics of

inadequate responders (Fletcher et al., 2011; Miciak et al., 2014). Agreement across single indicators of intervention response is limited (Barth et al., 2008; Fletcher et al., 2014). When using single indicators, group membership will fluctuate due to changes in criterion measures or testing occasion, an artifact of imperfectly reliable measures that inadequately measure the latent construct. Further, there is no “gold standard” for determining intervention response (Fletcher, Lyon, Fuchs, & Barnes, 2007). As a result, the selection of criterion measures and cut points is to some extent arbitrary. We therefore chose to employ response criteria that included: (1) multiple measures; (2) cut points fixed to the same benchmark sample to eliminate error introduced by different norming samples (Fletcher et al., 2014); and (3) multiple scores above the specified cut point in order to be considered an adequate responder. This third criterion reduces the number of false negatives (students considered to be adequate responders who continue to demonstrate reading risk and need interventions) at the expense of false positives (students considered to be inadequate responders who do not demonstrate subsequent difficulties in reading). This decision reflects our judgment that false positives are less deleterious than false negatives, given the relatively low expense and risk associated with continued intervention.

Benchmark sample.

The benchmark sample consists of all 846 students in grades 3 to 5 in the measurement study (Cirino et al., in press), but is not directly evaluated here. Rather, the benchmark sample is utilized to define the cut point for the bottom quartile on the criterion measures for intervention response. The measurement study oversampled struggling readers – a result of the planned overlap with the intervention study. To adjust scores to account for this oversampling of struggling readers, we weighted observations within the full sample of 846 to reflect a normal distribution of the GMRT within the population of the study (Cirino et al., in press, for additional details).

Measures

All measures were administered by trained examiners hired and supervised by the research team. Additional details regarding measures and procedures can be found at www.website.here.

Demographic Information.—Student demographic information was collected from the schools where the intervention took place. These data included student age, gender, free or reduced lunch status, special education eligibility, and race/ethnicity.

Executive Functioning Measures.—Details of the measures of EF can be found in Cirino et al. (in press). All EF measures were administered in fall, prior to the beginning of the reading intervention, except for the teacher questionnaires, which were administered in spring. Teacher questionnaires were administered in spring, after a full year in classrooms because a fall administration would have required teachers to answer questions about students with whom they are not familiar. For *Shifting*, parts of four subtests from the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001) were utilized (Design Fluency, Verbal Fluency, Color-Word Identification Test, and Trailmaking Test). Participants switched between two previously introduced actions or stimuli in all conditions.

Inhibition was assessed with two measures. A Go-No-Go task (Inquisit 3, 2003) required students to refrain from responding to an infrequent target. For Stop Signal (Inquisit 3, 2003), students tried to respond appropriately to two visual stimuli while inhibiting any response to an auditory stimulus. **Working memory as span/manipulation with planning (WMSMP)** included 5 measures, including the Listening Span subtest from the Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001), Corsi Blocks (Inquisit 3; 2003), two measures of the Tower of London test (Inquisit 3, 2003), and the WJ-III (Woodcock et al., 2001) Planning subtest. **Working memory as updating** used four N-back measures (Kirchner, 1958), also administered on the Inquisit platform. Using letters or complex (unnamable) shapes, students need to keep track of a running sequence and gauge whether a current stimulus matches one presented sometime previously; a d-prime measure was computed for each. **Generative fluency** used three subtests from the D-KEFS (Delis et al., 2001): Lexical Fluency, Categorical Fluency, and Design Fluency; for each subtest, students had one minute to generate as many exemplars as possible from a given letter (Lexical), semantic grouping (Category), or dot pattern (Design). **Self-regulated learning (SRL)** was assessed with the Contextual Learning Scale (Cirino et al., in press), a researcher-created self-rating scale that included new items, as well as existing or adopted items from the Motivated Strategies for Learning Questionnaire (Pintrich & DeGroot, 1990), School Motivation and Learning Strategies Inventory (Stroud & Reynolds, 2006), Child Behavior Rating Scale (Bronson et al., 1990), Motivations for Reading Questionnaire (Wigfield & Guthrie, 1997), Self-Regulation Strategy Inventory-Self Report (Cleary, 2006) and Patterns of Adoptive Learning Survey (Midgley et al., 1996). The scale assesses self-efficacy and effort, reading strategies, and perceived skill and preference for reading activities. Finally, **Metacognition** was assessed with the same-named scale of the Behavior Rating Inventory of Executive Function – Teacher (BRIEF; Gioia et al., 2000), and with the Inattention subscale from the Strengths and Weaknesses of ADHD symptoms and Normal behavior (SWAN; Swanson et al., 2006). Both questionnaires were administered in spring. Per the measurement study, factor scores were only derived for common EF, and specific EFs including: (1) working memory as span/manipulation with planning; (2) working memory as updating, (3) fluency, (4) self-regulated learning, and (5) metacognition.

Reading.—For descriptive purposes, we report scores in three areas of reading: decoding and spelling, fluency, and comprehension. All reading measures were administered in fall (pretest) and spring (posttest). For **Decoding and spelling**, the WJ-III Letter-Word Identification and Spelling subtests (Woodcock et al., 2001) were used to measure word reading accuracy and spelling. Test-retest reliability coefficients for children aged between 8–13 years range from .84 to .85. For **Fluency**, the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1998) Sight Word Efficiency subtest was used to measure word list fluency with real words and pseudowords. Alternate-forms reliability for this instrument exceeds .90 (Torgesen et al., 1998). For **Comprehension** the GMRT (MacGinitie et al., 2000) was administered in a timed, group setting. Alternate form reliability ranges from .80 to .87. The Passage Comprehension subtest from the WJ-III (Woodcock et al., 2001) is a cloze assessment in which students read a passage and fill in missing words. Test-retest reliabilities for children aged 8 to 13 range from .76 to .86.

Intervention

A full description of the intervention can be found in Vaughn et al. (2016). The standardized, researcher-delivered intervention met five times per week for approximately 16 weeks from November to April. Each session lasted approximately 30 minutes and was conducted with small groups of four to five students. The intervention consisted of several two-week thematic units of ten lessons each that ended with a Maze progress monitoring activity during the tenth lesson of the unit. During each session, students engaged in three activities: 1) word and concept building through the development of vocabulary knowledge (3–10 minutes), 2) text-based reading of narrative or expository passages (15–20 minutes), and 3) word study using decoding strategies (6–10 minutes).

Vocabulary.—Students received vocabulary instruction for 3–10 minutes during 9 out of each 10 sessions in a unit (6 explicitly taught vocabulary words) that pertained to the key social studies concepts presented in their readings. Words were taught using explicit instruction procedures, a sequence that included: (a) “kid friendly” definitions, (b) visual representations, (c) synonyms, (d) application of the word to personal experiences, and turn and talk procedures to utilize new words. Practice was distributed across the 10-day unit, and vocabulary learning was assessed utilizing a Maze activity to check for word understanding and inform instruction.

Text-based reading.—During the text-based reading portion of each unit, students read two types of texts: (a) stretch-texts or (b) fluency-texts. Stretch-texts included grade-level content, whereas the fluency-texts contained content at a student’s reading level. Stretch text instruction included extended text-reading and text-based questions that prompted students to integrate different sections of the text. During fluency-text instruction, students read passages from the QuickReads program (Hiebert, 2003). The instructional sequence included three steps: (1) text preview and questions; (2) an individual silent read; and (3) a summary activity, similar to stretch-text reading. To build fluency, students then engaged in repeated readings of the same text with an identified fluency focus (i.e., improved rate, improved accuracy, or better expression).

Word study.—Students practiced phonics skills through patterned word reading with words, phrases, and sentences. The word lists included multi-syllabic words, high-frequency words, and word patterns, depending on the abilities of the students.

Intervention implementation.—A team of 19 tutors who were hired and trained by the research team implemented the intervention. All tutors received ten hours of training on the implementation of the intervention as well as effective behavior management and engagement strategies. During the 16-week intervention, the tutors participated in 8 hours of continuing professional development, attended biweekly staff development meetings, and received coaching once every 2–3 weeks to maintain fidelity of implementation. The average total amount of instruction delivered to students at all sites was 42.2 hours (SD = 7.1 hours) over the course of the intervention.

Intervention fidelity.—Intervention implementation fidelity was evaluated using a four-point Likert-type rating scale for each component of the intervention. Global quality and fidelity scores were also rated on a four-point scale. Trained members of the research team, who had achieved interrater reliability > 90% completed all fidelity observations on a randomly selected sample of audio recordings of taught lessons. Across all components and tutors, the researcher-provided intervention received a mean implementation score of 3.71 ($SD = .24$, range 3.45 to 4.00), a mean global quality score of 3.71 ($SD = 0.50$, range 2.00 to 4.00), and a mean global fidelity score of 3.48 ($SD = 0.55$, range 2.00 to 3.82). Additional information pertaining to the intervention implementation and fidelity can be found in Vaughn et al. (2016).

Data Analysis

The pattern of EF performance over the one common and five specific EF factors across the three reader groups was assessed using a split-plot design to compare group performance across the six EF factors. We followed procedures outlined by Huberty and Olejnik (2006) for a descriptive discriminant analysis that allows for the interpretation of the contribution of different dependent variables to group discrimination. The advantage of this analysis plan is that it allows for simultaneous analysis of all variables and addresses the relative influence of the grouping variable (responder, Inadequate Responders, and Typical) on the outcome variables (the 6 EF factors). The design addresses two issues: (a) group differences across the set of EF factors, and (b) pattern or profile of group differences across the set of EF factors.

The analysis plan encompassed three steps. First, we evaluated the group-by-task interaction in the omnibus analysis to determine whether the effect of the grouping variable (responder status) was consistent across the set of dependent variables. In the presence of a statistically significant interaction, we proceeded with pairwise multivariate comparisons for each group combinations (typical versus adequate responder, typical versus inadequate responder, and adequate responder versus inadequate responder) to identify differences between groups. This analysis permits interpretation of specific groups differences on the set of dependent variables. To control for potential Type I error, a Bonferroni adjusted alpha of $p < .017$ (.05/3) was used for all pairwise multivariate comparisons. Each pairwise comparison computes a linear discriminant function, which maximally separates the groups. Following procedures described by Huberty and Olejnik (2006), we report three methods for interpreting the contribution of specific EF factors to the discriminant function: (1) canonical structure correlations; (2) standardized discriminant function coefficients; and (3) univariate contrasts. Univariate significance is evaluated at a Bonferroni adjusted alpha of $p < .008$ (.05/6). When conducting pairwise comparisons, univariate contrasts parallel the findings of canonical structure correlations. However, there is no statistical test of significance associated with the two multivariate methods for interpreting the contribution of specific variables to the discriminant function.

To address the value added contribution of EF for response, we followed procedures outlined by Stanovich and Siegel (1994), and used in later studies evaluating the cognitive attributes of adequate and inadequate responders to intensive interventions in early elementary

(Fletcher et al., 2011) and middle school (Miciak et al., 2014). Following these examples, we developed six regression models, one model predicting each of the EF factors included in this study. The three predictor variables comprise the two reading comprehension measures (WJ-III Passage Comprehension and GMRT) at pretest and a contrast reflecting adequate or inadequate responder status. The contrast determines whether there is unique variance associated with the relation between performance on the EF variable and responder status beyond the variance explained by performance on the reading measures at pretest. A statistically significant contrast for responder status would suggest that the continuum of severity hypothesis is insufficient to explain intervention responsiveness among this sample.

Results

Means and standard deviations on reading measures and on EF measures are provided in Table 2. A comparison of the *z*-score profiles for each group is presented in Figure 1. A split-plot design comparing the performance of the three groups on all six EF factors showed a significant group-by-task interaction, $F(10, 590) = 11.31, p < .0001, \eta^2 = .04$. To investigate this interaction, we performed three pairwise multivariate comparisons investigating main effects and interaction terms.

Discriminant Function

Typical versus adequate responder.—The interaction term for the comparison of the typical and adequate responder groups was statistically significant, $F(5,129) = 8.67, p < .001, \eta^2 = .04$. Table 3 shows that the common EF factor contributed most to the discriminant function maximally separating groups. The univariate contrast for the general EF factor was statistically significant, $F(1,133) = 44.47, p < .001, \eta^2 = .25$. No other univariate contrasts were statistically significant.

Typical versus inadequate responder.—The interaction term for the comparison of the typical and inadequate responder groups was statistically significant, $F(5,232) = 24.69, p < .001, \eta^2 = .05$. Table 3 shows that the common EF factor contributed most to the discriminant function maximally separating groups. The univariate contrast for the general EF factor was statistically significant, $F(1,236) = 45.52, p < .001, \eta^2 = .31$, as was the univariate contrast for BRIEF Metacognition, $F(1,236) = 7.75, p < .001, \eta^2 = .04$. No other univariate contrasts were statistically significant.

Adequate responder versus inadequate responder.—The interaction term for the comparison of the adequate responder and inadequate responder was not significant, $F(5,225) = 1.56, p > .05, \eta^2 = .00$. The main effect was significant, $F(5,225) = 33.16, p < .001, \eta^2 = .08$. Table 3 shows that the common EF factor contributed most to the discriminant function maximally separating groups. However, no univariate contrast was statistically significant.

Continuum of Severity

We created regression models predicting each of the six EF factors from this study. Each regression equation included three predictor variables: pretest performance on the two

criterion measures utilized to identify responder status (WJ-III Passage Comprehension and GMRT) and a dummy-coded grouping variable (1 for adequate responder and 0 for inadequate responders). The grouping variable was established to investigate whether there is unique variance in the EF factor associated with responder status beyond that accounted for by the reading variable. A significant contrast would suggest that reading performance at pretest is insufficient for predicting intervention responder status and that a continuum of severity hypothesis is insufficient to explain this pattern of results. In all six regression models, the contrast for responder status was not statistically significant, consistent with a continuum of severity hypothesis (beta coefficient range: -0.05 to 0.16).

Discussion

This study investigated whether EF (at pretest) is associated with responder status for a sample of 4th graders who are at significant risk for LD and who completed an intensive reading intervention. First, we evaluated to what extent performance on 6 (one common and five specific) EF factors was associated with group status: typical, adequate responder, and inadequate responder. Second, we evaluated whether initial EF performance explained unique variance in responder status beyond the variance explained in initial reading status. The present study extends previous literature in two ways. First, previous studies of EF have primarily utilized a correlational or cross-sectional design, or have investigated the relations of reading and EF in the context of direct EF training (typically with a focus on working memory). In contrast, the present study investigates relations of EF and reading in the context of an intensive reading intervention, ensuring that all participants have adequate instructional opportunity and permitting an investigation of whether initial EF might serve as a potential intervention adjuvant. Second, previous studies of the cognitive characteristics of inadequate responders have not included robust batteries evaluating EF (Stuebing et al., 2015). Given that inadequate response to evidence-based interventions is an inclusionary criterion for the identification of LD, it is important to better understand the performance of at-risk students who may differ on this dimension.

Differences in EF Attributes by Group

We found few statistically significant differences between the three groups across the 6 EF factors. As hypothesized, the largest separation occurred between the typical and inadequate responder groups. These differences were statistically significant in two domains: the common EF factor and the BRIEF/Metacognition factor. Correlations with the discriminant function indicated that group separation was driven largely by the common EF factor, with a smaller role for the BRIEF/Metacognition factor. Other correlations with the discriminant function were small and not statistically significant. That the largest separation occurred on the common EF factor is not surprising. In a separate study that also utilized this EF factor structure in the full sample of 846 across grades 3 through 5 in fall (but not addressing the intervention context), we found that the common EF factor demonstrated strong associations with “pretest” reading skills (Cirino et al., in review), the defining dimension of our group structure.

Given that groups were defined by reading comprehension performance, and the well-documented unconditional relations between reading comprehension and EF performance (Jacob & Parkinson, 2015), it is surprising that no statistically significant separation occurred between the inadequate responder and adequate responder groups, as hypothesized. Previous analyses of the cognitive attributes of inadequate responders have found statistically significant separation across a number of dimensions, including: phonological awareness; verbal knowledge, and rapid naming (Cho et al., 2015; Fletcher et al., 2011; Miciak et al., 2014). Additionally, reviews and meta-analyses have implicated several other cognitive dimensions, including: attention, problem behavior, nonverbal reasoning, and working memory (Al Otaiba & Fuchs, 2006; Nelson, Benner, & Gonzealez, 2003), although it should be noted that these studies were largely conducted with early elementary students. In explanation, we note that the effect size for the group-by-task interaction was of much smaller magnitude than in previous studies (e.g. Cho, et al.; Fletcher et al.; Miciak et al.), indicating that responder group status was more weakly associated with EF performance than previously investigated cognitive dimensions such as phonological awareness, rapid naming, or language. Given the well-documented and robust relations between reading comprehension and language processes (Catts, Hogan, & Adlof, 2005; Catts, Adlof, & Weismer, 2006), weaker associations for EF relative to language were expected. However, the failure to find any statistically significant between groups differences was not consistent with previous studies of this type.

The Unique Prediction of EF

Our second analysis investigated whether there was unique variance in responder status related to EF beyond that explained by reading performance at pretest. If EF and responder status were associated, even in a more comprehensive statistical model that included pretest reading performance, it would suggest that EF is uniquely predictive of intervention response. However, a lack of statistically significant, unique variance between EF and responder status beyond reading pretest suggests that observed differences in levels of EF are associated with the severity of reading difficulties, consistent with the continuum of severity hypothesis (Vellutino et al., 2006).

We in fact found no statistically significant associations between responder status and EF across the six contrasts. This result is consistent with previous studies of inadequate responders that have not found statistically significant “value added” for cognitive functioning above and beyond pretest for explaining responder status (Fletcher et al., 2011; Miciak et al., 2014). These results also align with the results of Stuebing et al., (2015) which documented small unique contributions for baseline cognitive characteristics above and beyond initial status in reading. Notably, Stuebing et al. included a wide range of potential predictors, including strong predictors of reading performance such as vocabulary and phonological awareness, but did not include EF as a predictor.

It is important to note that much of the work evaluating predictors and/or moderators of intervention response has focused on reading interventions (as in the present study), and often in early elementary grades. These results may not generalize to other grades and other academic areas (e.g. math). For example, there is intriguing evidence emerging in

mathematics research that working memory may moderate response to specific intervention protocols that differ in cognitive input and/or demands (e.g. Fuchs et al., 2013; Fuchs, et al., 2014; Swanson, Lussier, & Orosco, 2013). These studies suggest that a more granular understanding of the role of EF and academic achievement may be necessary.

EF, Reading Intervention, and Assessment

To understand the largely null results of the present study, we find it valuable to compare and contrast this study with our reading prediction study (Cirino et al., in review). We found moderate relations of EF with reading performance, even after controlling for a very robust set of covariates, including language and demographic variables. We consider this strong evidence for a unique role for EF in reading comprehension at this age range. These findings and this assertion may seem incongruent with the findings of the present study, although predicting responder status following an intensive intervention is a fundamentally different question than predicting reading performance at a single time point. The former question is a more difficult—though pragmatic—undertaking, because there is typically less variance associated with changes in reading performance in individuals with reading difficulties than with reading performance across a full range of achievement. Such analyses emanate from our experience in educational practice in which critical questions are shaped by considerations of who needs and who does not intensive intervention or who will continue to need interventions of even greater intensity and dosage after a robust intervention protocol, as documented in this study. Additionally, analyses of contributions of intervention response may point to promising avenues to improve intervention outcomes.

The results of the present study also have practical implications for the routine assessment of EF as part of typical school practice and for the identification of LD and/or attention problems. The present findings, particularly the findings of the second analysis which found little value-added for measures of EF, provide little support for the necessity of assessing EF as a potential predictor of response to intervention in reading. In other words, the assessment of EF does not provide important information that might be used to pre-select those with low EF for more intensive or differential types of intervention. In other forums, we have argued against the necessity of cognitive assessment for the identification of LD in reading for similar reasons (Fletcher & Miciak, 2017). The results from this study do little to dissuade us from such assertions related to EF and reading comprehension. However, we do not interpret the results of this study as demonstrating that EF play no role in reading comprehension, nor would we discourage future studies that attempt to leverage EF and EF-related skills to improve academic outcomes (e.g. Cirino et al., 2016; Peng & Fuchs, 2017). Additional work is needed in this area to determine whether and how best to incorporate EF instruction or academic instruction that explicitly incorporates procedures that scaffold for low EF, while maintaining an appropriate focus on the core instructional needs of struggling readers (i.e. explicit reading instruction).

Limitations

There are several potential explanations for the lack of statistically significant differences observed across both analyses. First, it is possible that the reading intervention incorporated in the present study was not ideally suited to leverage differences in EF at pretest to improve

intervention response. This study did not evaluate response to an intervention that specifically attempted to build EF skills in the context of reading intervention. While it included elements of self-regulation, including goal-setting and self-monitoring, these components did not constitute a large portion of instructional time, which was focused on building reading skills in core reading areas (e.g. decoding, fluency, and comprehension). Future studies should investigate whether an intervention that includes additional time instructing reading-related aspects of EF within the context of reading intervention might demonstrate more robust findings. Additionally, there are limitations related to the sample utilized in this study. The sample sizes for the adequate responder group and students identified as typical readers are relatively small. It is possible that with larger samples, more statistically significant differences would be observed.

Conclusion

We investigated whether performance on EF at pretest was predictive of responder status following an intensive reading intervention. At posttest, participants were classified into three groups (inadequate responders, adequate responders, and typical students) based on performance on two measures of reading comprehension. We then compared group performance on EF according to two analytic plans. Results indicated that students identified as typical readers significantly outperformed inadequate and adequate responders on a common EF factor, and significantly outperformed inadequate responders on the BRIEF/Metacognition factor, a teacher report measure. All other comparisons were not significant. Additional contrasts investigating whether EF predicted responder status above and beyond initial reading status were null, suggesting that observed mean differences in EF by group may reflect differences in reading ability. These results suggest that EF may not be a robust predictor of intervention response in reading comprehension. Thus, routine assessment of EF in school practice is currently not informative for predicting a students' likely response to intensive reading interventions.

Acknowledgments

This research was supported by Award Number P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to the University of Houston. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health.

References

- Al Otaiba S, & Fuchs D (2002). Characteristics of children who are unresponsive to early literacy intervention a review of the literature. *Remedial and Special Education*, 23(5), 300–316. doi: 10.1177/07419325020230050501
- Anderson VA, Anderson P, Northam E, Jacobs R, & Catroppa C (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, 20(1), 385–406. doi:10.1207/S15326942DN2001_5. [PubMed: 11827095]
- Barth AE, Stuebing KK, Anthony JL, Denton CA, Mathes PG, Fletcher JM, & Francis DJ (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences*, 18(3), 296–307. doi:10.1016/j.lindif.2008.04.004 [PubMed: 19081758]
- Best JR, Miller PH, & Jones LL (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, 29(3), 180–200. doi: 10.1016/j.dr.2009.05.002. [PubMed: 20161467]

- Bhattacharya A (2010). Children and adolescents from poverty and reading development: A research review. *Reading & Writing Quarterly*, 26(2), 115–139. doi:10.1080/10573560903547445
- Bierman KL, Nix RL, Greenberg MT, Blair C, & Domitrovich CE (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology*, 20(03), 821–843. doi:10.1017/S0954579408000394 [PubMed: 18606033]
- Blair C, & Razza RP (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78(2), 647–663. doi:10.1111/j.1467-8624.2007.01019.x [PubMed: 17381795]
- Bronson MB, Goodson BD, Layzer JI, & Love JM (1990). *Child behavior rating scale*. Cambridge: MA: Abt Associates.
- Cain K, Oakhill J, & Bryant P (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96(1), 31. doi:10.1037/0022-0663.96.1.31
- Catts HW, Adlof SM, Weismar SE (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49, 278–293. doi:10.1044/1092-4388(2006/023)
- Catts HW, Hogan TP, & Adlof SM (2005). Developmental changes in reading and reading disabilities. In Catts HW & A. G. (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cleary TJ (2006). The development and validation of the self-regulation strategy inventory— self-report. *Journal of school psychology*, 44(4), 307–322. doi:10.1016/j.jsp.2006.05.002
- Cirino PT, Ahmed Y, Miciak J, Taylor WP, Gerst E, & Barnes M (in press). A framework for executive function in the late elementary years, *Neuropsychology*.
- Cirino PT, Miciak J, Ahmed Y, Barnes MA, Taylor WP, & Gerst EH (in review). The Simple View and Executive Function: Association with Multiple Reading Skills.
- Connor CM, Morrison FJ, & Slominski L (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, 98(4), 665. doi:10.1037/0022-0663.98.4.665
- Delis D, Kaplan E, & Kramer J (2001). *Delis-Kaplan Executive Function Scale*. San Antonio, TX: The Psychological Corporation.
- Diamond A, Barnett WS, Thomas J, & Munro S (2007). Preschool program improves cognitive control. *Science*, 318(5855), 1387–1388. doi:10.1126/science.1151148 [PubMed: 18048670]
- Diamond A, & Lee K (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333(6045), 959–964. doi:10.1126/science.1204529 [PubMed: 21852486]
- Fitzpatrick C, & Pagani LS (2012). Toddler working memory skills predict kindergarten school readiness. *Intelligence*, 40(2), 205–212. doi:10.1016/j.intell.2011.11.007
- Fletcher JM, Lyon GR, Fuchs LS & Barnes MA (2007). *Learning disabilities: From identification to intervention*. New York, NY: Guilford Press.
- Fletcher JM, & Miciak J (2017). Comprehensive cognitive assessments are not necessary for the identification and treatment of learning disabilities. *Archives of Clinical Neuropsychology*, 32(1), 2–7. doi:10.1093/arclin/acw103 [PubMed: 27932345]
- Fletcher JM, Stuebing KK, Barth AE, Denton CA, Cirino PT, Francis DJ, & Vaughn S (2011). Cognitive correlates of inadequate response to reading intervention. *School Psychology Review*, 40(1), 3. [PubMed: 23125475]
- Fletcher JM, Stuebing KK, Barth AE, Miciak J, Francis DJ, & Denton CA (2014). Agreement and coverage of indicators of response to intervention: A multi-method comparison and simulation. *Topics in Language Disorders*, 34(1), 74. doi:10.1097/TLD.000000000000004 [PubMed: 25364090]
- Fletcher JM, & Vaughn S (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives*, 3(1), 30–37. doi:10.1111/j.1750-8606.2008.00072.x [PubMed: 21765862]
- Foorman BR, Schatschneider C, Eakin MN, Fletcher JM, Moats LC, & Francis DJ (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, 31(1), 1–29. doi:10.1016/j.cedpsych.2004.11.003

- Fuchs D, & Deshler DD (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice*, 22(2), 129–136. doi: 10.1111/j.1540-5826.2007.00237.x
- Fuchs L, & Fuchs D, & Vaughn S (2008). *Response to intervention: A framework for reading educators*. Newark, DE: International Reading Association.
- Gerst EH, Cirino PT, Fletcher JM, & Yoshida H (2015). Cognitive and behavioral rating measures of executive function as predictors of academic outcomes in children. *Child Neuropsychology*, 1–27. doi:10.1080/09297049.2015.1120860
- Gioia GA, Isquith PA, Guy SC, & Kenworthy L (2000). *Behavior Rating Inventory of Executive Function*. Odessa, FL: Psychological Assessment Resources, Inc.
- Goldstein K (1949). Frontal lobotomy and impairment of abstract attitude. *Journal of Nervous and Mental Disease*, 110(2), 93–111. [PubMed: 18136721]
- Gresham FM, Reschly DJ, Tilly WD, Fletcher J, Burns M, Christ T, ... & Shinn M (2005). Comprehensive evaluation of learning disabilities: A response to intervention perspective. *The School Psychologist*, 59(1), 26–30.
- Hale J, Alfonso V, Berninger V, Bracken B, Christo C, Clark E, ... & Dumont R (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly*, 33(3), 223–236. doi:10.1177/073194871003300310
- Harlow JM (1848). Passage of an iron rod through the head. *Boston Medical and Surgical Journal*, 39, 389–393.
- Hiebert EH (2003). *QuickReads – A research-based fluency program*. Parsippany, NJ: Pearson.
- Hofmann W, Schmeichel BJ, & Baddeley AD (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, 16(3), 174–180. doi:10.1016/j.tics.2012.01.006 [PubMed: 22336729]
- Huberty CJ, & Olejnik S (2006). *Applied discriminant analysis* (2nd ed.) New York, NY: Wiley.
- Huizinga M, Dolan CV, & van der Molen MW (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, 44(11), 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010. [PubMed: 16527316]
- Inquisit 3. (2003). Seattle, WA: Millisecond Software.
- Jacob R, & Parkinson J (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, 85(4), 512–552. doi:10.3102/0034654314561338
- Jimerson SR, Burns MK, & VanDerHeyden AM (2016). *Handbook of response to intervention*. New York, NY: Springer US.
- Jurado MB, & Rosselli M (2007). The Elusive Nature of Executive Functions: A Review of our Current Understanding. *Neuropsychology Review*, 17(3), 213–233. doi:10.1007/s11065-007-9040-z. [PubMed: 17786559]
- Kavale KA, Kauffman JM, Bachmeier RJ, & LeFever GB (2008). Response-to-intervention: Separating the rhetoric of self-congratulation from the reality of specific learning disability identification. *Learning Disability Quarterly*, 31(3), 135–150. doi:10.2307/25474644
- Kirchner WK (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 1–17. doi:10.1037/h0043688
- Kovaleski JF, VanDerHeyden AM, & Shapiro ES, (2013). *The RTI approach to evaluating learning disabilities*. New York, NY: The Guilford Press.
- Lee K, Bull R, & Ho RM (2013). Developmental changes in executive functioning. *Child Development*, 84(6), 1933–1953. doi:10.1111/cdev.12096. [PubMed: 23550969]
- Lesaux NK, Kieffer MJ, Kelley JG, & Harris JR (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, 51(6), 1159–1194. doi:10.3102/0002831214532165
- Lezak MD, Howieson DB, Bigler ED, & Tranel, 2012 *Neuropsychological assessment*. New York, NY: Oxford University Press.
- MacGinitie W, MacGinitie R, Maria K, Dreyer L, & Hughes K (2000). *Gates-MacGinitie Reading Tests-4*. Itasca, IL: Riverside.

- Melby-Lervåg M, & Hulme C (2013). Is working memory training effective? A meta-analytic review. *Developmental psychology*, 49(2), 270. doi:10.1037/a0028228 [PubMed: 22612437]
- Miciak J, Stuebing KK, Vaughn S, Roberts G, Barth AE, & Fletcher JM (2014). Cognitive attributes of adequate and inadequate responders to reading intervention in middle school. *School Psychology Review*, 43(4), 407–427. doi:10.17105/SPR-13-0052.1 [PubMed: 28579668]
- Midgley C, Maehr ML, Hicks L, Roeser R, Urdan T, Anderman E, ... & Middleton M (1996). Patterns of adaptive learning survey (PALS). Ann Arbor, MI: Center for Leadership and Learning.
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, 41(1), 49–100. doi:10.1006/cogp.1999.0734 [PubMed: 10945922]
- Molfese VJ, & Molfese DL (2002). Environmental and social influences on reading skills as indexed by brain and behavioral responses. *Annals of Dyslexia*, 52(1), 121–137. doi:10.1007/s11881-002-0009-6
- Morris RD, & Fletcher JM (1988). Classification in neuropsychology: A theoretical framework and research paradigm. *Journal of Clinical and Experimental Neuropsychology*, 10(5), 640–658. doi:10.1080/01688638808402801 [PubMed: 3066798]
- Nelson RJ, Benner GJ, & Gonzalez J (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research & Practice*, 18(4), 255–267. doi:10.1111/1540-5826.00080
- Peng P, & Fuchs D (2017). A randomized control trial of working memory training with and without strategy instruction: Effects on young children’s working memory and comprehension. *Journal of Learning Disabilities*, 50(1), 62–80. doi:10.1111/1540-5826.00080 [PubMed: 26156961]
- Piasta SB, Connor CM, Fishman BJ, & Morrison FJ (2009). Teachers’ knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading*, 13(3), 224–248. doi:10.1080/10888430902851364
- Pickering SJ, & Gathercole SE (2001). *Working Memory Test Battery for Children*. London: Psychological Corporation Europe.
- Pintrich PR, & De Groot EV (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33. doi:10.1037/0022-0663.82.1.33
- Pribram KH (1973). The primate frontal cortex – executive of the brain In Pribram KH and Luria AR (Eds), *Psychophysiology of the Frontal Lobes*. New York: Academic Press, pp. 293–314.
- Schatschneider C, Wagner RK, & Crawford EC (2008). The importance of measuring growth in response to intervention models: Testing a core assumption. *Learning and Individual Differences*, 18, 308–315. doi:10.1016/j.lindif.2008.04.005 [PubMed: 22224065]
- Stage SA, Abbott RD, Jenkins JR, & Berninger VW (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ—word reading discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities*, 36(1), 24–33. doi:10.1177/00222194030360010401 [PubMed: 15490889]
- Stanovich KE, & Siegel LS (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, 86, 24–53. doi:10.1037/0022-0663.86.1.24
- Stuebing KK, Barth AE, Trahan LH, Reddy RR, Miciak J, & Fletcher JM (2015). Are child cognitive characteristics strong predictors of responses to intervention? A meta-analysis. *Review of Educational Research*, 85(3), 395–429. doi:10.3102/0034654314555996 [PubMed: 26535015]
- Stroud KC, & Reynolds CR (2006). *The School Motivation and Learning Strategies Inventory*. Los Angeles, CA: Western Psychological Services.
- Stuss DT (1992). Biological and psychological development of executive functions. *Brain and Cognition*, 20(1), 8–23. doi:10.1016/0278-2626(92)90059-U [PubMed: 1389124]
- Suchy Y (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *Annals of Behavioral Medicine*, 37, 106–116. Doi:10.1007/s12160-009-9097-4 [PubMed: 19455377]

- Swanson J, Schuck S, Mann M, Carlson C, Hartman K, Sergeant J, & McCleary R (2006). Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN rating scales. University of California, Irvine Retrieved from: <http://vrosario.bol.ucla.edu/forms/SWAN.pdf>
- Tolar TD, Barth AE, Fletcher JM, Francis DJ, & Vaughn S (2014). Predicting reading outcomes with progress monitoring slopes among middle grade students. *Learning and Individual Differences*, 30, 46–57. doi:10.1016/j.lindif.2013.11.001 [PubMed: 24659899]
- Torgesen J, Wagner R, & Rashotte C (1998). *Test of Word Reading Efficiency*. Austin, TX: Pro-Ed.
- Torgesen JK, Wagner RK, Rashotte CA, Herron J, & Lindamood P (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia*, 60(1), 40–56. doi:10.1007/s11881-009-0032-y [PubMed: 20052566]
- VanDerHeyden AM, & Burns MK (2010). *Essentials of response to intervention* (Vol. 79). Hoboken, NJ: John Wiley & Sons.
- Vaughn S, Solís M, Miciak J, Taylor WP, & Fletcher JM (2016). Effects from a randomized control trial comparing researcher and school-implemented treatments with fourth graders with significant reading difficulties. *Journal of Research on Educational Effectiveness*, 9, 23–44. doi: 10.1080/19345747.2015.1126386 [PubMed: 28491206]
- Vellutino FR, Scanlon DM, Small S, & Fanuele DP (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of learning disabilities*, 39(2), 157–169. doi: 10.1177/00222194060390020401 [PubMed: 16583795]
- Wiebe SA, Sheffield T, Nelson JM, Clark CAC, Chevalier N, & Espy KA (2011). The structure of executive function in 3-year-old children. *Journal of Experimental Child Psychology*, 108(3), 436–452. doi:10.1016/j.jecp.2010.08.008 [PubMed: 20884004]
- Wigfield A, & Guthrie JT (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of educational psychology*, 89(3), 420. doi: 10.1037/0022-0663.89.3.420
- Woodcock RW, McGrew KS, & Mather N (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.
- Yeniad N, Malda M, Mesman J, van IJzendoorn MH, & Pieper S (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, 23, 1–9. doi:10.1016/j.lindif.2012.10.004

Figure 1.
EF z-score profiles by group

Table 1

Demographic data for the intervention and typical sample

Variable	Typical	Adequate Responders	Inadequate Responders
Age			
<i>M</i>	9.7	9.7	9.9
<i>SD</i>	0.38	0.44	0.58
% Male	43.7	57.8	54.5
% Economically Disadvantaged	88.5	95.0	93.7
% Limited English Proficient	0.0 ^a	37.5	59.3
Race or ethnicity, %			
African American	33.3	31.3	20.4
White	21.7	12.5	8.4
Hispanic	42.0	56.3	70.1
Other	2.9	0.0	0.6

^a Limited English Proficient (LEP) students were excluded from participation in the large measurement study and were therefore not available for comparisons as typically reading students. However, students classified as LEP were included in the intervention study and are included as at-risk students and classified as adequate and inadequate responders.

Table 2

Descriptive statistics for the intervention and typical sample at pretest and posttest

Measure	Inadequate Responders		Adequate Responders		Typical	
	<i>N</i> = 167				<i>N</i> = 71	
	<i>M</i>	<i>SD</i>			<i>M</i>	<i>SD</i>
<i>Reading at pretest</i>						
WJ-III Letter Word ID	87.02	11.21	93.64	7.16	107.30	10.40
TOWRE SWE	78.86	12.13	83.94	10.69	97.10	9.30
Gates MacGinitie RC	76.20	6.44	79.64	4.10	101.80	7.40
WJ-III Passage Comp	79.34	8.75	87.73	5.58	97.70	6.85
<i>Executive functions at pretest</i>						
Executive Functions	−0.69	0.64	−0.47	0.58	0.26	0.68
BRIEF/MetaCognition	−0.20	0.93	−0.16	0.81	0.20	0.92
Fluency	−0.18	0.75	−0.15	0.75	0.07	0.83
Self-Regulated Learning	0.04	0.95	0.03	0.94	0.12	0.90
Working Memory/Planning	−0.05	0.66	−0.05	0.67	−0.12	0.56
Working Memory/Updating	0.18	0.76	0.08	0.74	−0.01	0.62
<i>Criterion measures at posttest</i>						
Gates MacGinitie RC	80.88	6.54	92.03	4.44	106.18	10.29
WJ-III Passage Comp	79.87	8.12	92.09	3.77	99.62	6.88

Table 3

Canonical structure correlations and standardized coefficients for EF dimensions group status

	<u>typical versus adequate responder</u>		<u>typical versus inadequate responder</u>		<u>responder versus inadequate responder</u>	
	r_s	SDFC	r_s	SDFC	r_s	SDFC
Executive Functions	0.91 [*]	1.03	0.92 [*]	1.15	0.86	0.98
Fluency	0.25	0.29	0.25	0.13	0.11	0.05
WM/ Planning	-0.11	-0.07	-0.08	-0.25	0.00	-0.22
WM/ Updating	-0.12	-0.24	-0.2	-0.31	-0.34	-0.45
BRIEF MetaCognition	0.36	0.41	0.31 [*]	0.27	0.11	0.09
Self-Regulated Learning	0.09	-0.06	0.07	-0.07	-0.03	-0.18

Note: IR = Inadequate Responder, SDFC = standardized discriminant function coefficient.

^{*} Indicates univariate contrast significant at $p < .008$.