Assessing Teacher Understanding of Student Executive Functioning and Predictions to Academic Achievement

Jared B. Goldberg

University of Michigan

Mentor: Dr. Frederick Morrison

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science with Honors in Biopsychology, Cognition, and Neuroscience from the University of Michigan, 2018

Author Note

The author would like to thank to Sammy Ahmed and Nick Waters for their support during the duration of this project.

Abstract

This study examined elementary school teachers' judgments of their students' executive function (EF) skills. We collected surveys from teachers in four elementary schools in Southeastern Michigan. Questions adapted from the Children's Behavior Questionnaire (CBQ) along with original questions about working memory skills were included in the surveys. Using a structural equation model (SEM), we confirmed the internal consistency of teacher-reported EF measures, explored whether these measures were related to lab-based measures of EF, and tested if teacher or lab-based EF measures predicted academic achievement. Chi square test of model fit (χ^2 (223) = 168.1, p = .000), root mean square error of approximation (.05), Comparative Fit Index (.95), and Tuker-Lewis Index (.94) indicated excellent model fit. Survey responses strongly loaded onto three distinct latent factors, which were identified as attention control, response inhibition, and working memory. These factors also loaded strongly onto a hierarchically generated latent score of teacher-rated EF. Teacher-rated EF was significantly related to a latent lab-measured EF score. The lab-based EF score significantly predicted both math and literacy achievement, while the teacher-reported EF score did not. This final result contradicted our hypothesis, and discussion focused on potential reasons for this surprising pattern of results.

Keywords: executive function, academic achievement, teacher ratings, structural equation modeling

Assessing Teacher Understanding of Student Executive Functioning and Predictions to Academic Achievement

Executive function (EF) encompasses a variety of separate but related cognitive skills including self-control, working memory, and cognitive flexibility (Diamond, 2012). Executive processes develop throughout childhood and play a key role in cognitive functioning, behavioral and emotional control, and social competence (Anderson, Jacobs, & Anderson, 2010). These skills allow people of all ages to process information and adjust their behavior in ways appropriate for different situations. School is a social context where controlling behavior and cognitively manipulating information is key to success, particularly in the early schooling years. The ability to regulate behavior by leveraging EF has been shown to predict superior math and literacy outcomes in elementary-aged children (Blair & Razza, 2007).

Considering that EF is essential for academic success, it is important to understand how EF is structured throughout development. Cognitive batteries and surveys given to teachers or parents are two popular methods of measuring EF, and each has its strengths and weaknesses. Cognitive tasks have been shown to have strong ecological validity and are a proven indicator of achievement; however, these measures take much manpower and time to complete (Sadeh, Burns, & Sullivan, 2013). On the other hand, surveys are easier to administer and less time-consuming to collect; however, they can be systematically biased based on the respondent's attitudes and perceptions (Dekker, Ziermans, Spruijt, & Swaab, 2017). The purpose of this study was to test whether each of these measurements are associated and predict academic achievement, and to gauge the structure of EF in kindergarten students using surveys given to teachers.

Components of EF

Some argue that EF is a unified construct that manifests differently depending on context. Wiebe, Espy, and Charak (2007) tested this hypothesis with preschool children, and found that a single unified EF score best explained data from their executive control tasks. Others contend that EF is inherently fragmented and composed of unique processes. Miyake et al. (2000) answered this question using a college-aged sample and confirmatory factor analysis, and found that three specific cognitive domains were related but clearly separable. Although these and many other studies have examined the structure of EF using performance-based measures, much less is known about the structure of survey-rated EF.

One domain of EF that has been studied extensively is working memory. Working memory is the ability to monitor, store, and manipulate incoming information to complete cognitive tasks (Miyake et al., 2000). People of all ages utilize this function every day to answer questions, piece together different sources of information, and think critically. Working memory plays an important role in many aspects of thought, and has been demonstrated to predict children's academic achievement. For example, Bull & Scerif (2001) found that a higher working memory capacity is related to better mathematics outcomes for young children. Furthermore, children who display lower mathematical abilities were shown to perform poorly on working memory tasks. However, mathematical cognition is not the only skill affected by working memory. This EF domain has been shown to predict increased performance on standardized assessments of language comprehension (Cain, Bryant, & Oakhill, 2004). Considering that working memory appears to be key for both mathematical and literacy skills, it is important to monitor children's working memory abilities to understand how they can succeed in school.

Working memory is one of three important EF factors that we were interested in studying in the context of academic success. Attention control is another aspect of EF that affects children's ability to succeed in school. Attention control includes the capacity to selectively attend to specific stimuli and focus attention for long periods of time (Anderson, 2002). This EF domain appears to emerge in infancy and develop rapidly in early childhood (Anderson et al., 2010). The fact that attention control develops much earlier than other EF domains suggests that individual differences in this skill may explain variability in elementary school achievement. Previous studies have shown that attention control correlates with reading and mathematics outcomes across cultures (Lan, Legane, Ponitz, & Morrison, 2011). These results suggest the development of this skill influences each dimension of academic achievement.

Studies of attention control have been performed using many paradigms. Considering that attention-deficit/hyperactivity disorder (ADHD) involves the reduced ability to control one's focus on tasks and stimuli, one popular framework is to compare differences in EF skills between individuals with and without the disease. Groups with ADHD exhibited significant impairments in working memory and response inhibition tasks, and these effects were not explained by group differences in intelligence, academic achievement, or symptoms of other disorders (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Furthermore, previous research has found that stronger scores on attention-shifting measures in preschool are related to math and literacy ability in kindergarten (Blair & Razza, 2007). These studies reveal that the ability to focus one's attention is an important EF skill to develop for success in school.

Response inhibition is another essential domain within EF. This concept refers to the suppression of inappropriate behaviors that are no longer required, which supports flexible and goal-directed behavior in different environments (Verbruggen & Logan, 2008). In a school environment, the ability to control one's behavior is of utmost importance. Students are constantly expected to conduct themselves appropriately, especially when in front of their teachers, who have authority within the classroom context. Failure to do so consistently would greatly interrupt teaching routines. Response inhibition and inhibitory control are sometimes merged into one skill: effortful control. However, studies of EF structure have found that though related, each represents a unique construct (Miyake et al., 2000).

The relation between children's response inhibition skills and academic success has been shown across many disciplines. Inhibitory processes have been implicated in reading comprehension, vocabulary, and mathematics ability (St Clair-Thompson & Gathercole, 2007). Considering this prediction to numerous academic outcomes, response inhibition is considered a key EF domain for elementary school children. Outside of an experimental context, response inhibition can be gauged by susceptibility to distraction. Lesions of the prefrontal cortex, a brain region associated with executive skills, can produce impulsivity, distractibility, and deficits on EF tasks (Willcutt et al., 2005). This reveals that during the early schooling years, response inhibition along with both working memory and attention control is an important construct to measure when considering the relation between EF and academic achievement.

Behavioral Scales of EF

To conduct our study, we needed to identify survey-based tools that would allow us to gather accurate teacher-ratings of student EF. Numerous studies have leveraged parent and teacher-reported measures of EF, and many utilized the Behavioral Rating Inventory of Executive Function (BRIEF; Gioia, Isquite, Guy, & Kenworthy, 2000). The BRIEF scale is composed of 86 items that assess EF skills through gauging everyday behaviors and activities. Each BRIEF question has three possible responses related to the frequency of the given behavior

– never, sometimes, or often. The main advantage of this scale is the inclusion of reliable subscales reflecting specific EF sub-domains including inhibition, working memory, and emotional control. However, completion time and limited response options have caused some researchers to explore other options.

The Children's Behavior Questionnaire (CBQ) is another popular tool used to gain insight into children's executive skills, and similarly includes reliable subscales reflecting separate EF domains (Rothbart, Ahadi, Hershey, & Fisher, 2001). The CBQ was originally developed to provide a caregiver report of children's temperament, self-regulation, and reactivity. In the original CBQ, parents were asked to rate their child on a 7-point scale from "extremely untrue" to "extremely true" on nearly 200 items. The reliability of this scale was subsequently validated by numerous studies (Rothbart et al., 2001). Due to the length of the original CBQ, researchers wished to create more concise versions of the scale for use in different academic contexts. As a result, Short and Very Short Forms of the questionnaire were developed and rigorously tested for internal consistency, reliability, longitudinal stability, and correlation with the original CBQ (Putnam & Rothbart, 2006). We chose the CBQ due to its reduced length, high number of response variables, and previously successful implementation for domain-based EF studies.

Although these shorter versions exhibited lower internal consistency than the parent scale, the reliabilities of all but one scale – sadness – were greater than .70, widely considered the benchmark for sufficient internal consistency (Putnam & Rothbart, 2006; Nunnaly, 1978). This reveals that the Short and Very Short Forms of the CBQ are viable measures of children's temperament and cognitive skills. These revisions are useful because the original CBQ is time consuming, may require compensation to participants for completion, and includes measures that may not be within the scope of the research at hand. These revised scales have made it easier for researchers who lack the necessary resources to use the original CBQ. The CBQ has been further adapted to accommodate teacher-reported measures of children's abilities. Minor changes to the wording of questions were made to make items appropriate for the classroom context, but questions were not altered to the point that they changed the concept measured by the question (Schussler, 2012).

Teacher Ratings of EF

Evidence from lab-based measures suggests that EF is composed of multiple domains, with each contributing to academic success for elementary school children (Blair & Razza, 2007). However, less is known about teacher-rated EF, and whether it has a similar structure. Gioia, Isquith, Retzlaff, and Espy (2002) explored this question using the BRIEF scale with children in elementary school. Using confirmatory factor analysis, they found that a three-factor EF model offered the best fit. Factors incorporated questions geared towards behavioral regulation, emotional regulation, and metacognition. These findings suggest that teacher-rating scales reflect a component-based structure of EF.

Our understanding of whether teachers can gauge the cognitive abilities of their students, and whether these judgments also predict achievement in school, is still developing. Some studies suggest that teacher-ratings of EF skills correlate well with validated performance-based measures of the same skills. For example, Toplak, Bucciarelli, Jain, and Tannock (2008) examined the relationship between scores on individual EF domains covered in the BRIEF and cognitive measures of each respective domain. They found that teacher BRIEF ratings of response inhibition and working memory significantly correlated with cognitive measures of these same domains. Similarly, Shimoni, Engel-Yeger, and Tirosh (2012) found significant correlations between behavioral and cognitive measures of EF from children with ADHD.

Furthermore, many researchers have argued that behaviorally rated EF is a significant predictor of school achievement. Dekker et al. (2017) studied differences in behaviorally and cognitively measured EF and found that both concurrently explained emerging literacy skills. However, behaviorally rated EF did not explain any additional variance in math achievement above performance-based cognitive EF assessments for first and second grade students.

Gerst, Cirino, Fletcher, and Yoshida (2015) performed a similar study to examine the relation of behavioral and performance-based cognitive EF measures and their prediction of achievement. Similar to Dekker et al. (2017), this study found that behavioral ratings of EF did not explain any additional variance in math achievement above cognitive measures. On the other hand, this study suggested that cognitive measures did not explain any additional variance in literacy achievement above teacher-rated measures. Taken together, these studies offer support that teacher-rated EF can be a significant predictor of achievement alongside cognitive measures. This suggests both measures should be use to gain a full understanding of children's ability to succeed in school.

Other studies have cast doubt on the consistency between behavioral and cognitive EF measures. Toplak, West, and Stanovich (2013) reviewed 20 studies to understand the association between performance-based and behaviorally rated measures of EF. Their review found just 24% of correlations between these two types of EF measures to be significant. They also found the median correlation between associated measures to be .19, which signifies a relatively small linear relationship. Based on these results, the researchers concluded that performance and behavioral measures of EF do not assess identical cognitive constructs. Moreover, McAuley, Chen, Goos, Schachar, and Crosbie (2010) studied whether the BRIEF scale was associated with measures of behavioral disruption and performance-based measures of EF. They observed associations between BRIEF scores and behavioral impairment, but no significant associations between BRIEF scores and performance-based EF measures. Anderson, Anderson, Northam, Jacobs, and Mikiewicz (2002) performed a similar study examining the correlation between cognitive and BRIEF scores of EF; however, they used a sample of children with various brain diseases. They discovered few significant correlations between each corresponding measure, which suggested cognitive and behavioral measures gauge different constructs in this pathological context.

There is similarly a lack of consensus as to whether teacher-rated EF is a predictor of academic achievement. Sadeh et al. (2012) tested how well a behavioral EF scale developed from the Behavior Assessment System for Children (Reynolds & Kamphaus, 1992) predicted achievement. This scale was designed to allow teachers to identify underperforming students through low EF scores. Their results found scores on this teacher-rated EF task did not predict academic achievement for Kindergarten students.

Current Study

We were interested in performing our own analysis based on the lack of consensus in the literature regarding the structure of teacher-rated EF and its relation to lab-based EF and academic achievement. The current study had three aims. Our first aim was to assess the internal consistency of a teacher-reported EF measure. Our hypothesis was that teacher-rated EF would be composed of three unique sub-domains – working memory, response inhibition, and attention control. We tested this hypothesis using a confirmatory factor analysis to determine if teacher responses to survey questions about children's EF loaded on to separate latent variables. We subsequently created a latent overall teacher EF score. Our second aim was to explore whether our teacher EF measure corresponded with lab-based measures of these same skills. Our hypothesis was that teacher and lab-rated EF would be strongly associated. To test our second hypothesis, we created a latent lab-based EF score using established lab-based measures for working memory, attention control, and response inhibition, and assessed its relation to our teacher EF score. Our last aim was to test the relative prediction of our teacher-rated and lab-based EF to academic achievement. Our hypothesis was that both teacher-rated and lab-based EF would uniquely contribute to academic achievement scores. To test our final hypothesis, we compared the predictive validity of our latent lab and teacher EF scores to both literacy and mathematics achievement scores.

Method

Participants

Teachers and students were recruited from four elementary schools in Southeastern Michigan, as part of a larger ongoing study. This sample was socioeconomically and ethnically diverse. Across schools, the percentage of students receiving free or reduced lunch ranged from 2% to 72%, and the percentage of African American or Hispanic students ranged from 7.8% to 79.1%. Each teacher was asked to respond to a set of individualized questions regarding participating students from their class. In total, data were collected for 133 kindergarten students, of which 65 were males and 68 were females. Mean participant age at testing of academic achievement was 5.71 years, with a standard deviation of .37 years.

Survey Measures

Original questions as well as questions adapted from the CBQ were included in the survey. Our attention control battery consisted of five questions adapted from the Attentional

Focusing subscale of the CBQ Teacher Very Short Form. The response inhibition battery consisted of three questions adapted from the Inhibitory Control subscale of the CBQ Parent Long Form.

Because the CBQ scales did not include questions regarding children's working memory capacities, we created a scale of four questions tapping into skills related to working memory. The CBQ uses a seven-point scale, ranging from "extremely untrue" to "extremely true." However, in our survey we used a five-point response scale ranging from "never" to "always" to make judgments easier for teachers. All questions were recoded for positive valence, such that high scores reflected stronger EF skills. A full list of questions included and corresponding EF factors can be found in Table 1.

Lab Measures of EF and Academic Achievement

We utilized five lab-based measures of EF to compare with our teacher responses. Working memory was measured using the Backward Digit Span (DS-B), a subtest of the Wechsler Intelligence Scale (Wechsler, 1991). In this task, children are told sequences of numbers, and are asked to repeat the numbers in the reverse order. After each successful trial, the sequence increases in length by one number. There are two sections of testing, and each correct sequence earns one point on this assessment. Because children have to manipulate information stored in memory, the DS-B task is considered a measure of working memory, not short-term memory.

Response inhibition was assessed using the Head-to-Toes, Knees-to-Shoulders (HTKS) battery developed by Ponitz, McClelland, Jewkes, Connor, Farris, and Morrison (2008). In this game, children are told to do the opposite of what the experimenter has told them. In the initial task, subjects are asked to touch their head (or toes), and the correct response is to do the opposite. If subjects can correctly respond to this task, an advanced section is initiated where commands to touch the knees or shoulders are added. In this task, children are forced to inhibit their initial reaction to the experimenter's prompt as well as utilize working memory and focused attention to perform the correct response. The game-like nature of the assessment works well with young subjects, and analyses have proven its relevance for multiple EF functions (McClelland et al., 2014).

Attention control was assessed using standard scores of the Pair-Cancellation test (PC) taken from the Woodcock-Johnson III Tests of Cognitive Abilities (Woodcock, McGrew & Mather, 2001). In this test, children are presented with rows of pictures of dogs, balls, and cups. They are given three minutes to circle all pairs in which a picture of the ball is followed by a picture of the dog. Students are scored based on how many pairs they correctly recognized out of 69 possible correct answers.

Academic achievement in literacy and math was examined using Woodcock–Johnson III Tests of Achievement (Woodcock & Mather, 2000). To measure math achievement, we utilized standard scores of the Applied Problems subscale. This test requires children to answer questions by analyzing visual quantities, answering questions about money and time, and solving quantitative word problems. To measure literacy achievement, we utilized standard scores of the Letter-Word Identification subscale. This test requires children to name letters and words, with increasing difficulty.

Analytic Plan

As part of our SEM, a confirmatory factor analysis was conduced on eight items from the CBQ and three original questions. These variables were hypothesized to measure response inhibition, attention control, and working memory. More specifically, CBQ items 1, 2, 3, 4, and

5 were hypothesized to measure response inhibition, CBQ items 7, 9, and 12 were hypothesized to measure attention control, and original items 15, 16, and 17 were hypothesized to measure working memory. Maximum likelihood estimation was used over other methods for parameter calculation. Listwise deletion was used to handle missing values in the dataset. All statistical analyses were performed using Mplus statistical software version 6.11 (Muthén & Muthén, 2007).

Results

See Figure 1 for full SEM with all loadings and coefficients depicted graphically. A list of survey questions and their respective factors is reported in Table 1. Descriptive statistics for study variables are reported in Table 2. Variable correlations are reported in Table 3, and standardized model coefficients can be found in Table 4.

Results from the confirmatory factor analysis support the hypothesized three-domain structure (attention control, response inhibition, and working memory) of teacher-rated EF. Each assigned survey item hung together with their respective factors, each with a loading higher than .65, which exceeded the widely accepted cutoff of .40 (Stevens, 2001). Furthermore, the loadings of each latent teacher-rated EF factor onto a hierarchical latent score of overall teacher-rated EF each exceeded .85. All factor loadings for lab-based measures of working memory, response inhibition, and attention control onto a latent lab-based EF score exceeded .40. Chi square test of model fit (χ^2 (223) = 168.1, p = .000), root mean square error of approximation (.05), Comparative Fit Index (.95), and Tuker-Lewis Index (.94) indicated excellent model fit. No post-hoc model changes were made considering the good fit of the data with the model.

Results from the path analysis ($r = .50$, $p < .001$) suggest a significant positive correlation between the teacher-rated and lab-based EF scores. Our latent lab-based EF score significantly predicted literacy achievement ($\beta = .52$, $p < .001$). Similarly, the lab-based EF score significantly predicted math achievement ($\beta = .84$, $p < .001$). However, our latent overall teacher-rated EF score did not concurrently predict any variance in either math ($\beta = -.14$, $p > .05$) or literacy achievement ($\beta = -.10$, $p > .05$).

Discussion

We found that teacher-reported EF questions loaded strongly onto three distinct latent variables. Considering that the CBQ documentation specifically references which questions correspond to each EF domain and we developed questions pertaining to working memory, we were able to infer that these three latent factors represent response inhibition, attention control, and working memory. This result supported our hypothesis, and is notable considering the lack of consensus in the literature with regards to the construct of teacher-reported EF. Many studies have suggested a unitary latent structure of EF, especially in young children. Wiebe et al. (2007) ran an analysis using lab-based measures on preschool children with mean age 3.91, and found a single-factor model was sufficient to account for the data. Similarly, Welsh, Nix, Blair, Bierman, and Nelson (2010) found that the inherent structure of EF was sufficiently represented by a single latent factor for children three to five years old.

However, evidence suggests that the specific domains of EF begin to emerge as children grow older. Lerner and Lonigan (2014) studied the distinctness of EF skills in a preschool sample with mean age 4.65. The model with distinct factors for working memory and inhibitory control offered superior fit. Moreover, the correlation between each factor decreased with age. Furthermore, studies with adults have reveled a fractionated view of EF (Lamar, Zonderman, & Resnick, 2002). This further supports the idea that age contributes to the development of identifiably distinct EF domains. These findings suggest that distinct EF domains can emerge in children even prior to kindergarten, and that this separation increases with age. We were able to confirm a three-factor EF structure in our kindergarten sample, which supports the fact that these distinct cognitive domains arise behaviorally before or during children's time in kindergarten.

It is also important to note that nearly all of the studies examining factor structure of EF highlighted above were conducted using individualized cognitive measures. There is much less research examining the structure of EF as measured using surveys. Gioia et al. (2002) explored and confirmed a three-factor structure of the BRIEF, a survey-based measure of EF. They found latent scores drawn from responses on the BRIEF corresponded to general cognitive constructs including metacognition and emotional regulation. Our results, however, suggest the existence of distinct and more specific EF domains as opposed to generalized constructs. We identified response inhibition and attention control as distinct factors, which split the more general construct known as effortful control into two distinct EF dimensions. Our findings suggest that EF is composed of at least three distinct cognitive domains, and that these domains can be identified at a behavioral level with kindergarten-aged children.

Upon confirming our hypothesis of a three-factor EF structure as measured using surveys given to teachers, we were interested in creating a hierarchical latent variable representing overall teacher-rated EF. The loadings of our individual latent components on this overall score were strong. Furthermore, we loaded scores from lab-based assessments of working memory, attention control, and response inhibition onto a latent variable representing lab-measured EF. The loadings of these scores were also sufficiently strong. These results further support the idea that EF encompasses distinct but related cognitive skills (Diamond, 2012).

The second aim of our study was to compare lab and teacher-rated measures of EF, and our hypothesis was that these measures would be strongly associated. The results of our SEM supported this hypothesis and were in line with findings of previous studies. Toplak et al. (2008) found domain-specific relationships between scores on the BRIEF and corresponding cognitive measures. Furthermore, Shimoni et al. (2012) found significant correlations between behavioral and cognitively focused measures of EF in a sample of children with ADHD.

The fact that our teacher-reported measure of EF was significantly related to our lab-based measure suggests that teachers can successfully gauge each student's EF skills by observing their behavior. As a result, time-consuming lab-based batteries may not be needed for researchers to effectively understand individual and group differences in EF. This fundamental understanding of each student's cognitive strengths and weaknesses could allow teachers to tailor their instruction to benefit students who experience deficits in specific areas. Previous studies have revealed that low performers in elementary school made significant improvements in both cognitive functioning and academic achievement as duration of support increased (Campbell & Ramey, 1994). Our findings suggest teachers can accurately recognize differences in EF skills, so it is reasonable to assume they could intervene while students are at an earlier stage of development. This has the potential to greatly bolster later academic outcomes.

Our final aim was to examine the relative predictions of lab and teacher-rated EF to both math and literacy achievement. Our hypothesis was that teacher and lab-based EF scores would concurrently explain math and literacy achievement. Our results indicate that lab-based EF significantly predicts both mathematics and literacy achievement. However, we found that teacher-ratings of children's EF abilities showed no relation to either math or literacy achievement. This finding contradicted our hypothesis, and suggests that lab-based EF measures have higher predictive power to achievement compared to teacher-rated EF measures.

These findings contradict many previous studies. Dekker et al. (2007) found that both survey and performance-based measures of working memory and attention shifting concurrently explained differences in spelling achievement. Furthermore, a study by Gerst et al. (2015) suggested that cognitive measures of EF did not explain any additional variance in literacy achievement over and above teacher-rated measures. Though each of these previous findings did not suggest teacher-rated EF predicts math achievement, it is nonetheless noteworthy that in our study teacher-rated EF did not explain academic achievement in literacy either. These results suggest that teachers' behavioral ratings of students' EF do not robustly explain achievement, and lab-based measures remain essential for the prediction of school-based academic outcomes.

There are several potential explanations for why teacher-rated EF did not predict achievement in our model. It is possible that one teacher-rated EF domain is a strong predictor of achievement, whereas the others are not. If this were to be the case, these differential associations could cancel out in our aggregated latent teacher-reported EF variable. Furthermore, data collected from teacher surveys may be inherently biased. For example, teachers may have generally rated boys lower on a subset of questions than girls, a phenomenon that has been previously observed across cultures (Thorell, Veleiro, Siu, & Mohammadi, 2012). This or any similar systematic bias may explain why the survey-based measure was not predicting achievement. Also, both our lab-based EF battery and academic achievement measures were individualized tests assessed at the same time point. Teacher reports of children's EF skills, however, could be context-dependent. It may be the case that specific individualized measures more strongly predict other individualized measures. In this case, the classroom environment may have been a confounding factor. In future studies we hope to perform further analyses to reveal potential drivers of this lack of prediction.

The current study has many limitations, most notably a limited sample size. An increase in sample size will allow our conclusions to become more robust. A larger sample would also allow us to increase model complexity. For example, we would be further interested in comparing these teacher and lab-based EF ratings to applied behavioral measures of EF and neurological measures of EF. If teacher-rated EF similarly corresponded with these other measures, it would give further support to our conclusion that teacher can make accurate judgments about each student's inherent cognitive skills. This would give us a more complete understanding of differences and similarities between methods of measuring EF and its sub-domains.

Furthermore, our SEM was based on confirming our hypothesis of a three-factor structure of teacher-rated EF. We did not perform any comparisons of fit between models assuming either more or fewer EF factors based on our survey questions of interest. It is possible that we could have identified either a more or less complex teacher-rated EF structure by comparing many model iterations. This study serves as a snapshot of how teachers understand their students' EF skills. We would be further interested in exploring how the relationship between teacher and lab-rated EF and their relative prediction to achievement changes throughout development.

Conclusion

The present study suggests that EF in kindergarten-aged children, as measured by teacher-reports, is composed of at least three distinct sub-domains including working memory, response inhibition, and attention control. This finding lends support to the theory of a fragmented view of EF, and suggests that sub-domains emerge through perceptible behaviors even at this young age. We also found that that these teacher-reports of children's EF abilities are significantly associated with individualized lab-based ratings of EF. This finding is inconsistent with some previous research and suggests teachers can accurately assess the EF skills of their students simply by observing their behavior in the school setting. Although we cannot conclude that teacher and lab-based EF are measuring identical constructs, it appears that survey measures can be an effective means of identifying individual differences in EF. If such a survey measure were widely collected, teachers could leverage these judgments early on to identify students who appear to have deficiencies in key EF skills. Future research should focus on the most effective means of integrating these judgments with individualized interventions to bolster later behavioral and academic outcomes. Although teachers can identify individual differences in EF at a behavioral level, these findings suggest that these judgments may not be considered a robust predictor of academic achievement. Future studies should further explore relationships between EF measurements, and identify root causes of differential prediction to achievement.

References

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71-82. doi:10.1076/chin.8.2.71.8724

Anderson, V., Jacobs, R., & Anderson, P. J. (Eds.). (2010). Executive functions and the frontal lobes: a lifespan perspective. New York, NY: Psychology Press.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology*, *8*, 231–240. doi: 10.1076/chin.8.4.231.13509

Blair, C., & Razza, R. P. (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

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*(3), 273-293. doi: 10.1207/S15326942DN1903_3

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

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. *Child Development*, *65*(2), 684-698. doi: 10.2307/1131410

Dekker, M. C., Ziermans, T. B., Spruijt, A. M., and Swaab, H. (2017). Cognitive, parent and teacher rating measures of executive functioning: shared and unique influences on school achievement. *Frontiers in Psychology*, *8*(48). doi: 10.3389/fpsyg.2017.00048

Diamond, A. (2012). Executive functions. *Annual Review of Psychology*, *64*, 135-168.

Gerst, E. H., Cirino, P. T., Fletcher, J. M., & Yoshida, H. (2015). Cognitive and behavioral rating measures of executive function as predictors of academic outcomes in children. *Child Neuropsychology*, *23*(4), 381-407. doi:10.1080/09297049.2015.1120860

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. Odessa, FL: Psychological Assessment.

Gioia, G., Isquith, P. K., Retzlaff, P., & Espy, K. (2002). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF) in a clinical sample. *Child Neuropsychology*, *4*, 249–257. doi: 10.1076/chin.8.4.249.13513

Lamar, M., Zonderman, A. B., & Resnick, S. (2002). Contributions of specific cognitive processes to executive functioning in an aging population. *Neuropsychology*, *16*, 156-162.

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: a cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677-692. doi: 10.1016/j.jecp.2010.11.001

Lerner, M. D., & Lonigan, C. J. (2014). Executive function among preschool children: unitary versus distinct abilities. *Journal of Psychopathology and Behavioral Assessment*, *36*, 626-639. doi: 10.1007/s10862-014-9424-3

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society, 16,* 495–505. doi:10.1017/ S1355617710000093

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., & Pratt, M. E. (2014). Predictors of early growth in academic achievement: the head-toes-knees-shoulders task. *Frontiers in Psychology, 5*, 599. doi: 10.3389/fpsyg.2014. 00599

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Muthén, L. K., & Muthén, B. O. (2007). Mplus user's guide (6th ed.). Los Angeles, CA: Muthén & Muthén.

Nunnaly, J. (1978). Psychometric theory. New York: McGraw-Hill.

Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, *23*(2), 141-158. doi: 10.1016/j.ecresq.2007.01.004

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*, *87*(1), 102-112. doi: 10.1207/s15327752jpa8701_09

Reynolds, C. R., & Kamphaus, R. W. (1992). Behavior assessment system for children, manual. Circle Pines, MN: American Guidance Service, Inc.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at 3-7 years: the Children's Behavior Questionnaire. *Child Development, 72*, 1394-1408.

Sadeh, S., Burns, M., & Sullivan, A. (2012). Examining an executive function rating scale as a predictor of achievement in children at risk for behavior problems. *School Psychology Quarterly*, *27*(4), 236–246. doi:10.1037/spq0000012

Schussler, L. E. (2012). The development of the Children's Behavior Questionnaire, Teacher Version (CBQ-T). (Unpublished masters dissertation). University of Maryland, College Park, MD.

Shimoni, M., Engel-Yeger, B., & Tirosh, E. (2012). Executive dysfunctions among boys with attention deficit hyperactivity disorder (ADHD): performance-based test and parents report. *Research in Developmental Disabilities*, *33*(3), 858–865. doi: 10.1016/j.ridd.2011.12.014

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*(4), 745-759. doi: 10.1080/17470210500162854

Stevens, J. P. (2001). Applied multivariate statistics for the social sciences (4th ed.). Mahwah, Preschool Executive Control 31, NJ: Lawrence Erlbaum Associates.

Thorell, L. B., Veleiro, A., Siu, A. F. Y., & Mohammadi, H. (2012). Examining the relation between ratings of executive functioning and academic achievement: findings from a cross-cultural study. *Child Neuropsychology, 19*(6), 630-638. doi: 10.1080/09297049.09292012.09727792

Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2008). Executive functions: performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*(1), 53–72. doi:10.1080/09297040802070929

Toplak, M. E., West, R. F., and Stanovich, K. E. (2013). Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*, 131–143. doi: 10.1111/jcpp.12001

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*(11), 418-424. doi: 10.1016/j.tics.2008.07.005

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, *102*, 43–53. doi: 10.1037/a0016738

Weschler, D. (1991). Weschler intelligence scale for children. San Antonio, TX: Psychological Corporation.

Wiebe, S. A., Espy, K. A., & Charak, D. (2007). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*, 575–587. doi: 10.1037/0012-1649.44.2.575

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological Psychiatry*, *57*(11), 1336-1346. doi: 10.1016/j.biopsych.2005.02.006

Woodcock, R. W., & Mather, N. (2000). Woodcock–Johnson Psychoeducational Battery III. Itasca, IL: Riverside.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III tests of cognitive abilities. Itasca, IL: Riverside.

Table 1

Survey Questions and Corresponding EF Factors

Question	Label	Factor
When practicing an activity, has a hard time keeping her/his mind on it	CBQ 1	Attention Control
Will move from one task to another without completing any of them	CBQ 2	Attention Control
When drawing or coloring in a book, shows strong concentration	CBQ 3	Attention Control
When building or putting something together, becomes very involved in what s/he is doing, and works for long periods	CBQ 4	Attention Control
Is easily distracted when listening to a story	CBQ 5	Attention Control
Can wait before entering into new activities if s/he is asked to	CBQ 7	Response Inhibition
Has trouble sitting still when s/he is told to	CBQ 9	Response Inhibition
Can easily stop an activity when s/he is told "no"	CBQ 12	Response Inhibition
Loses track during complicated tasks and may eventually abandon these tasks	Q 15	Working Memory
Makes place-keeping errors (e.g., skipping or repeating steps)	Q 16	Working Memory
Shows incomplete recall of information	Q 17	Working Memory

Table 2
Descriptive Statistics for Study Variables

Variable	N	Range	Min	Max	Mean	SD
TQ_AC	133	16.80	12.60	29.40	23.11	4.15
TQ_RI	131	9.33	7	16.33	12.88	2.49
TQ_WM	132	9.33	7	16.33	12.66	2.70
PairCanc	110	48	75	123	104.94	9.62
HTKS	117	56	0	56	27.24	16.92
DS-B	117	7	0	7	2.03	1.59
Math	116	103	31	134	103.68	16.28
Literacy	117	91	67	158	113.58	13.07
Age at Testing	118	1.49	4.96	6.45	5.71	.37

Note: TQ_AC, TQ_RI, and TQ_WM refer to latent survey-based scores for attention control, response inhibition, and working memory.

Table 3

Pearson's Correlations for Study Variables

Variable	1	2	3	4	5	6	7	8
1. TQ_AC	-							
2. TQ_RI	.785**	-						
3. TQ_WM	.815**	.638**	-					
4. PairCanc	.209*	.017	.283**	-				
5. HTKS	.366**	.286**	.410**	.259**	-			
6. DS-B	.184	.042	.298**	.313**	.381**	-		
7. Math	.271**	.154	.349**	.354**	.433**	.524**	-	
8. Literacy	.206*	.035	.283**	.366**	.260**	.513**	.585**	-

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

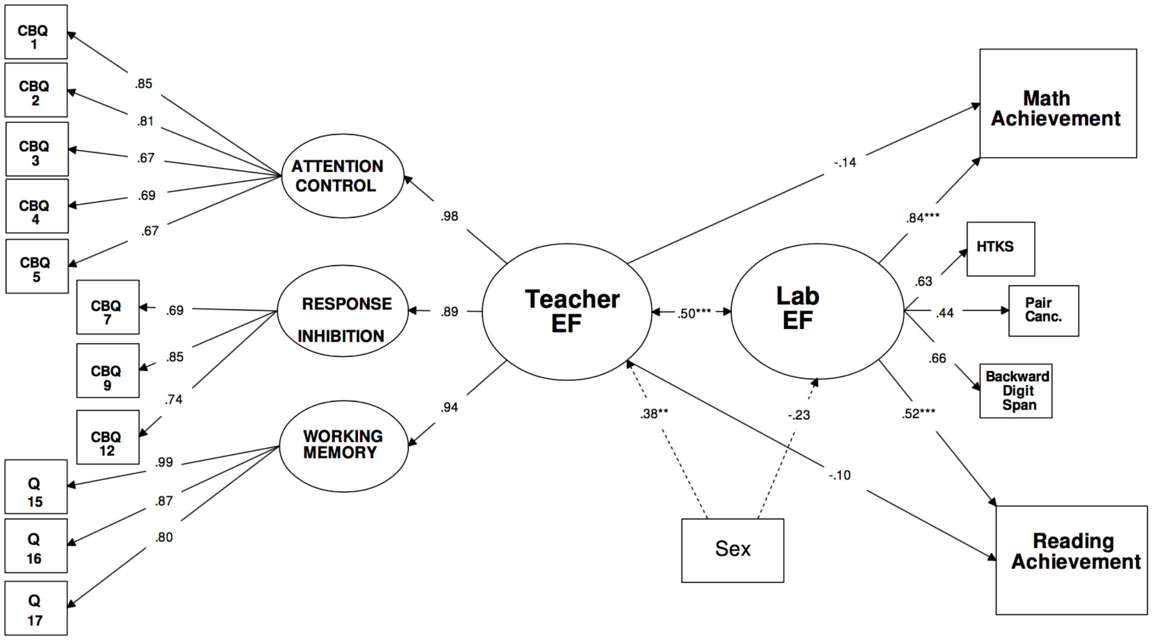
Note: TQ_AC, TQ_RI, and TQ_WM refer to latent survey-based scores for attention control, response inhibition, and working memory.

Table 4

Maximum Likelihood Estimates for a Three-Factor Model of Teacher-rated EF

	Estimate	S.E.	Est./S.E.	p-value
	Standardized Factor Loadings			
Attention Control				
CBQ 1	0.854	0.031	27.781	0.000
CBQ 2	0.811	0.037	22.163	0.000
CBQ 3	0.666	0.055	12.079	0.000
CBQ 4	0.691	0.052	13.267	0.000
CBQ 5	0.672	0.055	12.270	0.000
Response Inhibition				
CBQ 7	0.685	0.058	11.767	0.000
CBQ 9	0.845	0.039	21.939	0.000
CBQ 12	0.741	0.052	14.176	0.000
Working Memory				
Q 15	0.988	0.014	68.564	0.000
Q 16	0.867	0.026	33.439	0.000
Q 17	0.800	0.036	22.167	0.000

Note: CBQ refers to the Children's Behavior Questionnaire. All standardized factor loadings were statistically significant at $p < .01$.



Model fit statistics: χ2 (223) = 168.1, p = .000; CFI = .95; TLI = .94; RMSEA = .05.
**p* < .05. ***p* < .01. ****p* < .001.

Figure 1. SEM model for teacher-rated and lab-based measures of EF and predictions to academic achievement. Circles represent latent variables, while squares represent observed variables.