
Christina Weiland\textsuperscript{a,*}, M. Clara Barata\textsuperscript{b} and Hirokazu Yoshikawa\textsuperscript{c}

\textsuperscript{a}University of Michigan, Ann Arbor, MI, USA
\textsuperscript{b}Instituto Universitário de Lisboa (ISCTE-IUL), Cis-IUL, Lisboa, Portugal
\textsuperscript{c}New York University, New York, NY, USA

Despite consensus in the developmental literature regarding the role of executive function (EF) skills in supporting the development of language skills during the preschool years, we know relatively little about the associations between EF skills, including all EF components, and vocabulary skills among preschool-aged children. In this paper, we address this gap by fitting structural equation models to examine the associations between beginning-of-preschool EF and end-of-preschool receptive vocabulary, as well as the association between beginning-of-preschool receptive vocabulary and end-of-preschool EF skills, to data from 400 preschool-aged children. We find that EF skills at the beginning of preschool are a significant predictor of receptive vocabulary skills at the end of preschool, controlling for receptive vocabulary at the beginning of preschool. However, receptive vocabulary skills at the beginning of preschool do not predict EF skills at the end of preschool, controlling for EF at the beginning of preschool. Our findings are suggestive of the important role EF plays in supporting growth in receptive vocabulary skills, above and beyond earlier vocabulary levels, and of the direction of the developmental pathway between EF and receptive vocabulary in the preschool years. The discussion highlights the need for further research in

*Correspondence to: Christina Weiland, School of Education, University of Michigan, 610 E. University Avenue, Ann Arbor, MI 48104, USA. E-mail: weilandc@umich.edu
Executive function (EF) skills are defined as a set of cognitive processes integral to the emerging self-regulation of behaviour and the development of social and cognitive competence in young children (Blair, 2002). The development of EF skills mirrors the growth of the prefrontal cortex during the early years (Blair, 2002; Posner & Rothbart, 2000). That is, EF skills start developing in children as early as the end of the first year (Rueda, Posner, & Rothbart, 2005), followed by substantial developmental growth in EF during the preschool years (Welsh, Nix, Blair, Bierman, & Nelson, 2010; Zelazo & Müller, 2002). Recent findings suggest that EF skills predict children’s short-term and longer-term developmental outcomes. Children with higher EF skills are more likely to outperform their peers (with lower EF) in their academic skills in preschool and elementary school (Blair & Razza, 2007; Howse, Lange, Farran, & Boyles, 2003; McClelland, Acock, & Morrison, 2006; McClelland, Morrison, & Holmes, 2000; McClelland et al., 2007; Welsh, Nix, Blair, Bierman, & Nelson, 2010), and in their college outcomes (McClelland, Acock, Piccinin, Rhea, & Stallings, 2013). Children with higher levels of EF during their early years also become more socially competent in preschool and elementary school (Carlson & Moses, 2001; Espy, Sheffield, Wiebe, Clark, & Moehr, 2011; Hughes & Ensor, 2008, 2011), and seem to have better health, wealth, and crime outcomes in adulthood (Moffitt et al., 2011).

Scholars of EF argue that preschool tasks often demand the use of three EF constructs or factors: inhibitory control, attention shifting, and working memory (Best & Miller, 2010; Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Friedman et al., 2006). Inhibitory control is defined as the ability to suppress prepotent responses. Children exhibiting inhibitory control can stop automatic responses and behaviours and use more alternative, less automatic ones. In the classroom, they can block distractions from peers and focus on learning (Carlson & Moses, 2001; Diamond, Carlson, & Beck, 2005). Attention shifting is the ability to direct one’s attention as necessary to a given stimulus (Morrison, Ponitz, & McClelland, 2010). This factor is sometimes also called cognitive flexibility (Best & Miller, 2010). Children with well-developed attention-shifting skills are able to sustain and switch attention from one stimulus to the next and shift tasks when necessary (Bierman et al., 2008). Conceivably, in the classroom, this ability allows the child to navigate transitions between different classroom activities successfully. Working memory allows preschoolers to retain and manipulate unrelated information for learning (e.g. Morton & Munakata, 2002). For example, in preschool, children are often expected to recall, apply, and associate new knowledge in different classroom activities.

There is now growing consensus from studies of preschoolers in the US, UK, and Chile that EF skills in preschool are unidimensional in nature (Barata, 2011; Fuhs & Day, 2011; Hughes, Ensor, Wilson, & Graham, 2010; Wiebe, Espy, & Charak, 2008). Differentiation into more distinct factors appears to occur later in development (Friedman et al., 2006; Friedman et al., 2008; Huizinga, Dolan, & van der Molen, 2006; Miyake et al., 2000).

There is increased interest in understanding the relation between EF skills and other developmentally important skills, including children’s receptive vocabulary. Receptive vocabulary is essential in communicating with others and has been consistently linked with later reading outcomes (Scarborough, 2001). A recent
meta-analysis found that across studies, oral language skills in the preschool years—defined as the ability to produce or comprehend spoken language, including vocabulary and grammar—were predictive of decoding and spelling skills in first and second grades (National Early Literacy Panel, 2008). Vocabulary is also strongly tied to the emergence of phonological awareness (Cooper, Roth, Speece, & Schatschneider, 2002; Lonigan, 2006; Scarborough, 2001). Vocabulary acquisition sets the stage for children’s implicit comparisons between similar-sounding words, and such comparisons are one basis for the emergence of phonological awareness (Goswami, 2001). Phonological awareness in turn also has been linked to later literacy outcomes, including decoding, spelling, and comprehension skills in early elementary school (National Early Literacy Panel, 2008; Scarborough, 2001).

At present, the direction of the developmental pathways between EF skills and language skills is unclear. Some evidence suggests that EF skills in preschool may support development of language skills. For example, empirically, higher EF skills in preschool are positively associated with better language skills in both preschool and elementary school (Gathercole & Pickering, 2000; Gathercole, Pickering, Knight, & Stegmann, 2004; Howse et al., 2003; Lehto, 1995; McClelland et al., 2000, 2007; Ponitz, McClelland, Matthews, & Morrison, 2009). Conceptually, EF skills help children focus on multiple streams of information at the same time, monitor errors, and make decisions in light of available information, which is essential in children’s acquisition and development of initial language skills (Diamond, 2013). Neurobiological (Blair, 2002; Blair & Diamond, 2008; Blair et al., 2011; Blair & Peters, 2003) and cognitive mechanisms (Best and Miller, 2010; Diamond, 2013) have been purported for this pathway. According to the neurobiological models, early experiences determine stress physiology and cortex development (including the development of the prefrontal cortex) that are likely to regulate EF skills in the organization of information in goal-directed activities, such as learning (Blair et al., 2011). According to the cognitive model, EF refers to a family of top–down mental processes needed to control one’s actions and thoughts at will in order to learn from one’s surroundings, particularly when relying on automatic processes would be ill-advised or impossible (Diamond, 2013).

However, it is also possible that this developmental pathway could be reversed; language skills in preschool could support the acquisition and development of EF skills. One hypothesized mechanism for this association is that better receptive vocabulary could build EF skills by enhancing children’s outer and then inner speech. Improved self-talk may then improve EF, as children become better able to plan and monitor their behaviour (Fuhs & Day, 2011; Zakin, 2007). This mechanism is aligned with Vygotskian theories that posit that private speech in young children is a precursor of verbal thinking and serves as a carrier of thought at the time when most higher mental functions, such as EF skills, are not fully developed (Bryck & Fisher, 2012; Vygotsky, 1978). Private speech also helps children regulate both their overt and mental behaviours. In fact, recent empirical evidence seems to support the role of self-verbalization in the performance of demanding EF situations like task switching (Bryck & Mayr, 2005; Emerson & Miyake, 2003). Empirically, several studies have found that higher verbal ability is positively associated with children’s performance in EF tasks among toddlers and preschoolers (Carlson, Mandell, & Williams, 2004; Carlson, Moses, & Claxton, 2004; Hughes & Ensor, 2007; Kirkham, Cruess, & Diamond, 2003; Miller & Marcovitch, 2011).

To our knowledge, only one prior study has examined the direction of the developmental pathways between EF and language abilities in preschool-aged children. Among a sample of 132 children who ranged in age from three to five
and who spoke English as a first language, Fuhs and Day (2011) found that fall verbal ability supported spring EF, controlling for fall EF. There was no reciprocal relation between the two domains; that is, fall EF did not support spring verbal ability, controlling for fall verbal ability.

THE PRESENT STUDY

There is a need for additional studies of the developmental pathways between language skills and EF that posit and fit latent factor models to investigate reciprocal associations between language skills and EF skills, as such models account for the error in each manifest variable (Kline, 2011) and allow the researcher to test hypotheses about the latent structure of EF directly and to examine its dimensionality (Barata, 2011). In that vein, we investigated the structural relation between a latent factor representation of EF at the beginning of preschool and subsequent receptive vocabulary skills at the end of preschool using crossed-lagged structural equation models. In accordance with the hypothesis that there may be a reciprocal association between EF skills and receptive vocabulary, we also examined the structural relation between beginning-of-preschool receptive vocabulary and a latent factor representation of end-of-preschool EF. This model is presented in Figure 1. Specifically, we addressed the following two research questions:

1. Is beginning-of-preschool EF, as represented by a latent factor structure, positively associated with later receptive vocabulary skills, controlling for prior receptive vocabulary skills?
2. Is beginning-of-preschool receptive vocabulary positively associated with later EF skills, as represented by a latent factor structure, controlling for prior EF skills?

Figure 1. Structural model describing the fitted associations between latent factor structures for executive function and receptive vocabulary at the beginning and end of preschool.

Key: BDS₁ = Backward Digit Span at Time 1; PT₁ = Pencil Tapping at Time 1; FDS₁ = Forward Digit Span at Time 1; DCCS₁ = Dimensional Change Card Sort at Time 1; ATT₁ = Assessor's Ratings of Attention Shifting at Time 1; BDS₂ = Backward Digit Span at Time 2; PT₂ = Pencil Tapping at Time 2; FDS₂ = Forward Digit Span at Time 2; Vocab₁ = PPVT at Time 1; Vocab₂ = PPVT at Time 2.
METHOD

Participants and Setting
The study sample included 400 children attending an urban public preschool programme in 2009–2010 in a large urban public school district in the Northeastern USA. The preschool programme is open to any 4-year-old in the city; there are no income requirements or other criteria for enrollment as there are in many preschool programmes (Barnett et al., 2010). The programme has been shown to have moderate effects on children’s receptive vocabulary and small effects on children’s EF skills (Weiland & Yoshikawa, 2013). Sample children represent a subset of the 1049 children who agreed to participate in an evaluation of the 4-year-old preschool programme in fall 2009 (Weiland & Yoshikawa, 2013). In spring 2010, 400 children tested in fall 2009 were randomly selected to be tested again, as part of the district’s regular bi-annual programme monitoring. On average, there were 178 days between the fall and spring testing sessions (range = 61–259 days).

Participants were diverse in their background characteristics and were approximately representative of the district’s 4-year-old population. Overall, 67% of sample children received free/reduced lunch, 12% were classified as having a special need, and 51% were male. In terms of language, 27% spoke Spanish at home, 52% spoke English at home, and the remaining 21% spoke a language other than Spanish or English at home. In terms of racial/ethnic diversity, 44% were Hispanic, 29% were Black, 17% were White, and 11% were Asian. Thirteen children of mixed race/ethnicity who were tested in both fall and spring were excluded from the present study because of the unclear meaning/definition of that subgroup. Average child age at time of fall testing was 4.5 years and at the time of spring testing, 5.0 years.

All study teachers had at least a bachelor’s degree. In the study year, a literacy curriculum, Opening the World of Learning (OWL; Schickedanz & Dickinson, 2005), and a mathematics curriculum, Building Blocks (Clements & Sarama, 2007a), were in place system-wide in 4-year-old classrooms and had been in place since the 2007–2008 school year. Both curricula have shown positive results in other contexts (Ashe, Reed, Dickinson, Morse, & Wilson, 2009; Clements & Sarama, 2007b), although evidence for the effectiveness of the OWL is mixed (Dickinson, Freiberg, & Barnes, 2011; Dickinson, Kaiser, et al., 2011). While the OWL directly targets vocabulary, neither the OWL nor Building Blocks specifically targets EF skills.

Procedures
Child assessments were conducted in one-on-one pull-out sessions in the child’s school with trained child assessors. In the fall session, child assessors had to prove reliability on the battery of tests and show good rapport/child management skills in both simulated and real testing situations. Children were assessed on 10 tests, five of which we use in the present study (Peabody Picture Vocabulary Test (PPVT), Pencil Tap, Forward Digit Span (FDS), Backward Digit Span (BDS), and the Dimension Change Card Sort; see Weiland & Yoshikawa, 2013, for the full list of fall assessments). On average, the battery of tests took approximately 45–50 min to administer. Assessors were instructed to test children in one session if possible but to divide the session into smaller segments if children showed signs of fatigue. Because of the session length, we randomly varied the order of the tests to limit the possibility of systematically biasing results because of child fatigue. The assessors visited classrooms in fall 2009, as close to the start of the school year as teacher/school schedules and study staffing would allow.
A different study team assessed study children in spring 2010. Spring assessors also had to prove reliability on the battery of tests and show good rapport/child management skills in both simulated and real testing situations. Assessment sessions were similar to the fall sessions in that the order of the tests was systematically varied. Because of funding reasons, the spring battery was considerably shorter, taking on average 20–30 min, and assessors administered four direct child assessments (PPVT, Pencil Tap, FDS, and BDS). Because EF is a unidimensional construct in early childhood and to keep the testing session no longer than 20–30 min, we elected not to administer the Dimension Change Card Sort in the spring, as it was the most time-intensive measure of EF in our fall battery.

In both testing sessions, all children were tested in English. Also, in both sessions, assessor reliability was assessed using study-developed checklists for assessor testing and results recording. Assessors had to exhibit use of the correct rules of the assessments (start/stop rules, item order, etc.), as well as reliable recording of participants’ responses.

**Measures**

We describe the instruments that we used to create the manifest variables that are incorporated into the hypothesized structural models organized by the common latent construct from which they are hypothesized to be derived.

**Working memory**

In both the fall and the spring, assessors administered two tests of children’s working memory—FDS (Gathercole & Pickering, 2000) and BDS (Gathercole & Pickering, 2000). In both tasks, the assessor read aloud a string of numbers to the test child, with approximately a 1-s pause between digits. The child then either had to repeat back exactly what the assessor said (in FDS) or reverse the string of numbers (in BDS). Before trial items were administered, the child had to pass a practice trial, demonstrating that he or she understood the directions of the task. In both tasks, the child had two opportunities to demonstrate that he or she had a working digit span memory of a given number of digits. For the FDS, there were 10 possible items—two items with two numbers, two items with three numbers, and so on, up to two items with six digits. For the BDS, there were eight possible items; the two most difficult items had five digits. The tests were discontinued when a child missed two items in a row. FDS was scored from 1 to 6, while BDS was scored from 1 to 5. The score represented the child’s digit span memory (i.e. a 2 represents a digit span memory of two digits). Because both BDS and FDS scores were not normally distributed in our sample, we treated them as ordinal variables in all confirmatory factor analysis (CFA) and structural equation modelling (SEM) analysis.

The BDS and FDS measure different dimensions of working memory. BDS is considered a measure of the central executive component, while FDS is considered a measure of phonological loop. These are but a subset of the factors underlying the common working memory latent factor (Alloway et al., 2004). The structure of working memory in prekindergarten children is outside of the scope of the present study.

**Attention shifting**

In the fall, children were assessed with one measure of attention shifting, the Dimensional Change Card Sort (DCCS). In the DCCS (Frye, Zelazo, & Palfai, 2000),
1995; Zelazo, 2006), children were shown target cards that varied along the dimensions of colour and shape (e.g. red and blue, and rabbits and boats). Children learned to sort the cards according to one dimension (shape or colour), and then they were asked to sort the cards according to the other dimension. After up to four practice trials on each of the two dimensions to confirm that children understood the rules, the assessor administered up to 10 trials on the DCCS. After six trials, if a child had missed more than one trial, the test was discontinued. If the child had missed only one or zero trials, the assessor continued until trial 10. To ease demand on children’s working memory and to help minimize risk of conflating measurement of the working memory and attention-shifting constructs, between each trial, children were either asked to restate the rules or were reminded of the rules. The DCCS final total score represented the number of trials (out of 10) in which the child managed to shift attention from the prior criterion to correctly sort the cards according to the new criterion. The total score was non-normally distributed, with considerable pileup of observations at the lowest and highest possible values at each time point, as described in previous studies using this measure (Bierman et al., 2008). As such, we created an ordinal version of the total score, where DCCS = 2 if the total score was equal to 10 (a perfect score), DCCS = 1 if the total score less than 10 but greater than 0, and DCCS = 0 if the total score was equal to 0. In all CFA and SEM models, we treated DCCS as ordinal.

Inhibitory cognitive control

One measure was used to assess inhibitory control in the fall and spring: the Pencil Tapping (PT; Diamond & Taylor, 1996) task. In this cognitive performance task, the child was asked to tap twice if the evaluator tapped once and tap once if the evaluator tapped twice. Assessors first administered a set of practice trials (up to six) to ensure that children understood the rules of the task. Children who passed the practice were then given 16 total trials. The task measures mostly children’s inhibitory cognitive control and, to a lesser degree, working memory and fine motor activity (Bierman et al., 2008). Scores represented the correct number of trials out of 16. The score variable was non-normally distributed, so we created a discretized version where PT = 3 if the score was greater than 14, PT = 2 if the score was between 7 and 14, PT = 1 if the score was between 1 and 6, and PT = 0 if the score was 0. In all CFA and SEM models, we treated PT as ordinal.

Receptive vocabulary

Children’s receptive vocabulary was measured using the PPVT-III (Dunn & Dunn, 1997), a nationally normed measure that has been widely used in diverse samples of young children (Love et al., 2005; Wong et al., 2007). The test has excellent split-half and test–retest reliability, as well as strong qualitative and quantitative validity properties. It requires children to choose which of four pictures best represents a stimulus word. In analysis, we used the raw score from the PPVT-III, as the standardized score reduced the variance in our sample.

Data Analytic Approach

We used SEM with Mplus 6.0 (Muthen & Muthen, 2010) Los Angeles, CA, United States, to answer our research questions. Models with CFI and TLI values above .9 and RMSEA values below .05 were considered to have ‘good’ fit (Kline, 2011). We began by first using CFA to evaluate the psychometric structure of the EF construct. Next, we fit a longitudinal configural invariance measurement model across the two
time points (Little, Preacher, Selig, & Card, 2007). We then examined longitudinal metric invariance (equal factor loadings across time). Our invariance work was essential to determine whether EF skills have the same factor structure at the beginning and end of preschool (Kline, 2011). Because we treated the manifest indicators of EF at time 1 and time 2 as ordinal, we computed \( \chi^2 \) difference tests using the DIFFTEST option in Mplus, as standard \( \chi^2 \) difference testing is inappropriate when using WLSMV estimation. In all longitudinal models, EF measures gathered at time 1 and time 2 were allowed to load unto a shared latent measurement factor (e.g. FDS1 and FDS2 loaded onto a shared latent factor, and BDS1 and BDS2 loaded onto a different shared latent factor). This was mathematically equivalent to allowing an EF measure collected at time 1 to correlate with the same measure collected at time 2 (Kline, 2011).

Finally, we added receptive vocabulary and fit a structural model that addressed our research questions. The structural model included a clustering adjustment for the nesting of students within school, as well as a set of control covariates (student age, a set of dichotomous indicators for Black, Hispanic, and Asian race/ethnic groups, and a set of dichotomous indicators for Spanish home language and non-Spanish and non-English home language, for student free-lunch and reduced-lunch status, and for student gender). We allowed the covariates to correlate with EF and receptive vocabulary at time 1 and estimated structural paths between the covariates and the time 2 EF and receptive vocabulary variables. In all models, we assumed data were missing at random, and we used full information maximum likelihood to impute missing data (Graham, 2009).

RESULTS

To examine patterns among the EF skills and receptive vocabulary outcomes, we first computed sample descriptive statistics and zero-order correlations (see Table 1). As shown in Table 1, on all child outcome measures, the sample average was higher at time 2 than time 1. Sample correlations revealed highly statistically significant relations between all variables (\( p < .001 \); see Table 2). Correlations between EF assessments were generally moderate in magnitude (.40 or lower). Correlations among all variables were relatively stable from the beginning to the end of prekindergarten. Missing data information on EF and receptive vocabulary manifest indicators are displayed in Table 1. There were no missing data on student-level covariates.

Confirmatory factor analysis of the hypothesized latent EF construct at time 1 had adequate fit (\( \chi^2=2.10, \text{CFI}=1.00, \text{TLI}=.99, \text{RMSEA}=.01 \)). At time 2, there were only three indicators of EF, so the CFA model was just identified and fit was perfect (\( \chi^2=.00, \text{CFI}=1.00, \text{TLI}=1.00, \text{RMSEA}=.00 \)). In a longitudinal configural invariance model in which we fit time 1 and time 2 latent EF models, we found evidence for configural invariance (\( \chi^2=12.24, \text{CFI}=.99, \text{TLI}=.99, \text{RMSEA}=.02 \)). Likewise, we found that longitudinal metric invariance, in which we constrained the factor loadings at time 1 and time 2 to be equal, held. Comparing the metric invariance to the configural invariance models, \( \chi^2 \) difference tests indicated that we failed to reject the null hypothesis that the metric invariance model had worse fit than the configural invariance model (\( \chi^2(2, N=17)=3.50, p=.17 \)).

Figure 1 displays the fitted structural equation model that addresses our research questions. As shown in the figure, EF at time 1 was strongly correlated with receptive vocabulary at time 1 (\( r=.65, p<.001 \)), while EF at time 2 showed only a small correlation with receptive vocabulary at time 2 (\( r=.33, p<.001 \)). EF at time 1 was a strong predictor of EF at time 2 (\( \beta_{22}=.65; p<.001 \)), while vocabulary
Table 1. Descriptive statistics on the manifest indicators of executive function skills and receptive vocabulary at the beginning of preschool (time 1) and end of preschool (time 2)

<table>
<thead>
<tr>
<th>Construct</th>
<th>Assessment</th>
<th>Time 1</th>
<th>Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N (% missing)</td>
<td>Mean</td>
</tr>
<tr>
<td>Working memory</td>
<td>FDS</td>
<td>374 (6)</td>
<td>3.85</td>
</tr>
<tr>
<td>Working memory</td>
<td>BDS</td>
<td>351 (12)</td>
<td>1.26</td>
</tr>
<tr>
<td>Inhibitory control</td>
<td>PT</td>
<td>363 (9)</td>
<td>1.50</td>
</tr>
<tr>
<td>Inhibitory control</td>
<td>DCCS</td>
<td>379 (5)</td>
<td>1.01</td>
</tr>
<tr>
<td>Receptive vocabulary</td>
<td>PPVT</td>
<td>378 (6)</td>
<td>46.80</td>
</tr>
</tbody>
</table>

Note: PT and DCCS were coded as ordinal versions of the raw totals. BDS, Backward Digit Span; DCCS, Dimensional Change Card Sort; FDS, Forward Digit Span; PPVT, Peabody Picture Vocabulary Test; PT, Pencil Tapping.
at time 1 was a moderate predictor of vocabulary at time 2 ($\beta_{11} = .41; p < .001$). Also shown in Figure 1, EF at time 1 was positively associated with receptive vocabulary at time 2, controlling for receptive vocabulary at time 1 ($\beta_{12} = .32; p < .001$). Thus, with regard to our first research question, EF at the beginning of preschool was a statistically significant predictor of end-of-preschool receptive vocabulary, controlling for beginning-of-preschool receptive vocabulary.

Regarding our second research question, receptive vocabulary at the beginning of preschool (time 1) was not associated with EF at the end of preschool (time 2), after controlling for EF at time 1. The estimated structural parameter for this association was non-significant ($p > .05$) and close to zero in magnitude.

Finally, we examined the sensitivity of our conclusions to our decision to treat EF manifest variables as ordinal. Specifically, we refit our final model using continuous versions of the DCCS and PT tests and treating the two digit-span tests as continuous. Results were robust to treatment of manifest EF indicators as ordinal or continuous (e.g. the magnitude of the association between receptive vocabulary at time 1 and EF at time 2 was 0.04 ($p = .75$), while the association between EF at time 1 and receptive vocabulary was 0.35 ($p < .001$)). We also checked the sensitivity of our results to the coding of the BDS. That is, many children did not pass practice trials on the BDS ($N = 252$ in the fall and $N = 263$ in the spring) in which they were asked to reverse in order two given numbers. Accordingly, these children were assigned a score of 1, representing a backward digit span of one digit. We examined response patterns for these children, and we set to missing scores for children who repeated the assessor on the practice items (e.g. repeated ‘3, 6’ both times; $N = 75$ children in the fall and $N = 55$ children in the spring) and imputed their scores, as these children may possibly have not understood the task.
posed to them. Results of refitting the final model using the recoded BDS variable were not meaningfully different than those in which we used with the originally coded variable; the magnitude of the association between receptive vocabulary at time 1 and EF at time 2 was 0.01 (\(p = .87\)), while the association between EF at time 1 and receptive vocabulary was 0.32 (\(p < .001\)).

**DISCUSSION**

We found that EF skills at the beginning of preschool were positively associated with end-of-kindergarten receptive vocabulary, controlling for beginning-of-preschool receptive vocabulary. By using a psychometrically valid latent representation of EF that included all EF components at the beginning of preschool and at the end of preschool with longitudinal invariance, we provide evidence of the role of EF skills in supporting growth in language skills.

We also found that children’s EF skills increased from the beginning of preschool to the end of preschool, which is to be expected given the rapid development of EF in the preschool years. Interestingly, vocabulary in the beginning of preschool was less predictive of vocabulary in the end of preschool, compared with the association between EF in the beginning of preschool and EF in the end of preschool, when we controlled for each of these skills in the beginning of preschool. These findings are consistent with the curricula children were exposed to; that is, the preschool programme directly targeted vocabulary but not EF. Children’s receptive vocabulary scores are accordingly less stable over time, as this preschool programme has shown greater success in improving children’s vocabulary than children’s EF (Weiland & Yoshikawa, 2013).

We also found that receptive vocabulary skills at the beginning of preschool did not predict end-of-preschool EF skills, controlling for beginning-of-preschool EF skills. This result is consistent with Welsh et al. (2010), who used a composite of attention and working memory test scores and found that beginning-of-preschool emergent literacy skills did not predict end-of-preschool EF skills. Ours is the first evidence to our knowledge of a similar null finding for a beginning-of-preschool receptive vocabulary skills and end-of-preschool EF skills link, with all three subdomains of EF included. Interestingly, also similar to our results, Welsh et al. (2010) found that children’s beginning-of-preschool EF skills did predict their end-of-preschool emergent literacy. The consistency of our results with Welsh et al. (2010) thus raises interesting questions regarding the possibility that EF plays a similar supporting role in the development of both language and literacy skills among preschool-aged children. Vocabulary and emergent literacy skills are distinct but related domains, and empirical evidence suggests that vocabulary is positively and moderately correlated with emergent literacy skills such as letter identification and early writing (Barata, 2011).

Our null finding, however, conflicts with that of Fuhs and Day (2011) who found support for an association between fall verbal ability and spring EF but not between fall EF and spring verbal ability. While it is not clear why the results of their study and ours diverge, we note that their latent measure of verbal ability included one manifest indicator of receptive vocabulary and two manifest indicators of expressive vocabulary. It is possible that the association between verbal ability and EF is different for receptive versus expressive vocabulary. The hypothesized mechanism—that better vocabulary supports improved inner speech, which may improve EF (Fuhs & Day, 2011; Zakin, 2007)—may only hold for expressive vocabulary or general verbal ability. Also, the Fuhs and Day study...
did not include all three hypothesized dimensions of EF, while ours did. Further, although we controlled for home language, it is possible that our results differ because our sample included children whose home language was not English, while the Fuhs and Day study did not. Finally, it is not clear in the Fuhs and Day (2011) study whether the Head Start intervention (a preschool programme primarily targeted to low-income children) they studied directly targeted language or EF skills, while the public preschool we studied directly targeted language skills. While not directly comparable, public preschool programmes have been found to have larger impacts on language than Head Start programmes (National Forum on Early Childhood Policies and Programs, 2010; U.S. Department of Health and Human Services, Administration for Children and Families, 2010). While neither our study nor the Fuhs and Day study is an intervention study, the developmental pathways between language and EF may be contingent upon contextual factors and may vary depending on the skills targeted within a particular intervention and by their effectiveness in doing so. The differences in the findings of our studies highlight the need for thorough documentation of the specifics of the developmental context and the importance of conducting such studies in different kinds of preschool settings given that the majority of 4-year-old children in developed countries attend preschool (OECD, 2011).

In terms of the mechanism that may underlie the developmental pathway between EF and receptive vocabulary, our study cannot distinguish between the two possible explanations discussed early in this paper—a neurobiological (Blair, 2002; Blair & Diamond, 2008; Blair et al., 2011; Blair & Peters, 2003) or cognitive mechanism (Best and Miller, 2010; Diamond, 2013). Ultimately, both are viable explanations that are consistent with our results regarding the positive relationship between beginning-of-preschool EF and end-of-preschool receptive vocabulary. Our findings do, however, offer counterevidence to the proposed Vygotskyian-based hypothesis that improved inner speech may be enough to support the development of children’s self-regulatory skills, at least when receptive vocabulary is used as the proxy for inner speech and the setting is language-rich.

Our study has several important limitations. First, because of the lack of measures that are truly equatable for English and non-English speaking children and because sample children spoke a diversity of languages, all children were tested in English. Also, our study did not include a measure of IQ, so we cannot completely eliminate the hypothesis that the relation found between EF and language is a function of intelligence. Measures of fluid intelligence are in fact highly correlated with independent measures of EF (Duncan et al., 2008). However, the fact that we did not find a relation between vocabulary and EF makes this hypothesis less likely, although we cannot rule it out entirely. A third limitation is that all sample children were enrolled in a preschool programme that used curricula that targeted their vocabulary skills. Accordingly, we cannot disentangle the effects of setting from natural development. However, given that the majority of children in developed countries attend preschool (OECD, 2011), this is a relatively minor limitation of our study, as preschool represents an increasingly common context of development for children. In addition, even though we used a latent representation of EF, it is important to note that children had to exercise language skills to complete each task. This task impurity problem affects all studies of children’s EF and is further compounded by the fact that the construct validity for most EF tasks has not been well established (Espy & Kaufmann, 2002; Miyake et al., 2000). The relatively strong correlation between EF and receptive vocabulary at time 1 in our study ($r = .66; p < .001$) may reflect this problem; previous researchers have noted that this problem complicates the interpretation of the
patterns of association between composites of EF measures and emergent academic skills (e.g. Espy et al., 2004).

Future research that can identify the mechanisms underlying the developmental pathways between EF and vocabulary among preschool children would be a welcome addition to the literature. In addition, we know little about the optimal way to promote children’s EF skills in preschool, and possibilities include classroom-based interventions (Bierman, Domitrovich, et al., 2008; Bierman et al., 2008; Diamond et al., 2007; Domitrovich, Cortes, & Greenberg, 2007; Lipina & Colombo, 2009; Raver et al., 2008, 2011; Webster-Stratton, Reid, & Stoolmiler, 2008; Weiland & Yoshikawa, 2013) or computerized training (Dowsett & Livesey, 2000; Klingberg, Forssberg, & Westerberg, 2002; Klingberg et al., 2005; Olesen, Westerberg, & Klingberg, 2004; Rueda, Rothbart, McCandliss, Saccomano, & Posner, 2005; Thorell, Lindqvist, Nutley, Bohlin, & Klingberg, 2009). While the evidence base for the former is comparably stronger than that of the latter, rigorous evaluations of classroom-based versus computerized EF interventions for preschoolers in diverse settings are needed to more definitely inform specific programmatic choices in the preschool period and to translate developmental studies like ours into practice.

Note

1. As shown in Figure 1 and as described earlier in this paper, the DCCS test was not given at time 2. Essentially, we treated this time 2 indicator as missing at random in our CFA models. Because EF is generally thought to be a unitary construct in preschool-aged children and because we confirmed that configural and metric invariance held across time, the missingness of this indicator at time 2 affects the reliability but not the substantive meaning of the EF construct at time 2.

ACKNOWLEDGEMENTS

Christina Weiland and Hirokazu Yoshikawa’s work on this study was funded by the Institute for Education Sciences. Special thanks to participating families, teachers, principals, early childhood coaches, district staff, Carolyn Lazyer and colleagues at Abt Associates, and the Wellesley Centers for Women.

REFERENCES


