

The Lexical Stroop Sort (LSS) picture-word task: A computerized task for assessing the relationship between language and executive functioning in school-aged children

Makeba Parramore Wilbourn · Laura E. Kurtz ·
Vrinda Kalia

Published online: 19 August 2011
© Psychonomic Society, Inc. 2011

Abstract The relationship between language development and executive function (EF) in children is not well understood. The Lexical Stroop Sort (LSS) task is a computerized EF task created for the purpose of examining the relationship between school-aged children's oral language development and EF. To validate this new measure, a diverse sample of school-aged children completed standardized oral language assessments, the LSS task, and the widely used Dimensional Change Card Sort (DCCS; Zelazo, 2006) task. Both EF tasks require children to sort stimuli into categories based on predetermined rules. While the DCCS largely relies on visual stimuli, the LSS employs children's phonological loop to access their semantic knowledge base. Accuracy and reaction times were recorded for both tasks. Children's scores on the LSS task were correlated with their scores on the DCCS task, and a similar pattern of relationships emerged between children's vocabulary and the two EF tasks, thus providing convergent validity for the LSS. However, children's phonological awareness was associated with their scores on the LSS, but not with those on the DCCS. In addition, a mediation model was used to elucidate the predictive relationship between phonological awareness and children's performance on the LSS task, with children's vocabulary fully mediating this relationship. The use of this newly created and validated LSS task with different populations, such as preschoolers and bilinguals, is also discussed.

Keywords Executive functions · Language · Phonological processing · Receptive vocabulary · Phonological awareness · Working memory

Executive function (EF) is an umbrella term that refers to the cognitive processes involved in the conscious control of thoughts and action (Anderson, 2002; Posner & Rothbart, 2000). Although EF has been extensively studied in adult populations, there has been less research on the development of EF in children (Lan, Legare, Ponitz, Li, & Morrison, 2011). Researchers often ascribe certain types of cognitive processes as fundamental components of EF, such as selective attention, inhibitory control, cognitive flexibility, working memory, and particular information-processing skills (e.g., fluency and processing speed) that are utilized for problem solving and goal-directed activities (Blair & Razza, 2007; Müller, Zelazo, Lurye, & Liebermann, 2008; Willoughby, Blair, Wirth, & Greenberg, 2010). Many of these cognitive processes have been shown to play a vital role in emotion regulation, school readiness, and academic achievement in both typical and atypical populations (Blair & Razza, 2007; Espy, Bull, Martin, & Stroup, 2006; McClelland et al., 2007a). As a result, the exploration of developmental changes in executive functioning has become a specific focus of recent empirical research.

The development of EF is ongoing through childhood and into early adulthood; however, the developmental trajectory remains unclear (Anderson, 2002). Much of the research investigating the developmental trajectory of the abilities underlying EF has been focused on pinpointing the development of memory (Baddeley, 1986), attention (Lan et al., 2011), inhibition (Carlson, Moses, & Breton, 2002; Carlson, Moses, & Hix, 1998; Frye, Zelazo, & Burack, 1998), or language (Kirkham, Cruess, & Diamond, 2003; Luo, Luk, & Bialystok, 2010). The findings have been mixed, however, resulting in little agreement as to which processes are most

M. P. Wilbourn (✉) · L. E. Kurtz
Department of Psychology and Neuroscience, Duke University,
417 Chapel Drive, 208 Soc-Psych Building,
Durham, North Carolina 27708, USA
e-mail: makeba.wilbourn@duke.edu

V. Kalia
Department of Psychology, Worcester State University,
Worcester, MA, USA

influential in the development of EF (Baddeley, 1986; Carlson et al., 2002; Carlson et al., 1998; Frye et al., 1998; Kirkham et al., 2003; Marcovitch & Zelazo, 2009).

Although it is generally acknowledged that some relationship exists between language development and EF, the exact dimensions of this relationship are less understood (Singer & Bashir, 1999). It has previously been found that children with higher EF skills tend to have larger vocabularies (McClelland et al., 2007b). As a result, some researchers have speculated that the cognitive abilities that underlie EF are mediated by language (Denckla, 1996; Jacques & Zelazo, 2001). In order to use cognitive processes for goal-directed activities, children must use language to regulate their thoughts and behavior. Hence, empirical studies have focused on the regulatory role of language in the development of various cognitive processes or executive functions (Jacques & Zelazo, 2001; Kirkham et al., 2003; Müller et al., 2008; Vygotsky, 1986). For example, Frye et al. (1998) contended that as children develop cognitively, they gain the ability to maintain more complexity in the rules they use to solve problems. Marcovitch & Zelazo (2009, p. 8) proposed that one manner in which children gain this ability to formulate rules is through self-directed speech (e.g., “If I use the bathroom, then I get to pick a sticker”). This internal self-directed speech may allow children to maintain rules in working memory, which serves as a running reminder for children to exploit when solving problems and making decisions (Marcovitch & Zelazo, 2009). As a result, Marcovitch and Zelazo have argued that children’s developing language abilities are a driving force behind their developing executive functioning skills.

However, research also indicates that the relationship between EF and language is not merely defined by children’s use of internal speech to reiterate rules that serve a regulatory purpose. Previous work with children has shown that verbal labels can create a semantic link between children’s knowledge base and their current experience (e.g., Homer & Nelson, 2009). For example, Kirkham et al. (2003) contended that during an EF task, children’s use of verbal labels may serve to provide a scaffold to achieving an abstract representation. In Kirkham et al.’s study, the authors encouraged children to verbally label objects (e.g., “What is this?”) and the relationships of the objects (e.g., “Where does this go?”) in the task. The finding was that labeling helped children achieve a more complex abstract representation, which in turn redirected their attention to the relevant aspects of the task (Kirkham et al., 2003). The findings from similar studies have suggested that verbal labeling (e.g., a child naming the box in which the relevant stimuli are hidden) is particularly unique, in that it allows children to hold two mental representations (e.g., the word and the representation of the word) in their mind simultaneously, ultimately permitting deeper processing of

the information (Miller & Marcovitch, 2011). This implies that children’s working memory capacity, specifically the phonological loop, may be implicated in the relationship between EF and language development.

Baddeley and colleagues (Baddeley, 1996; Baddeley & Hitch, 1974), in their influential model of working memory, proposed that the central executive, responsible for allocating attention and processing information, is influenced by two essential supporting systems: the visuo-spatial sketchpad and the phonological loop. The visuo-spatial sketchpad operates on spatial and visual information, whereas the phonological loop operates on sound-based or auditory information. Recent research has shown that children’s visual spatial skills might be influenced by exposure to orthography in literacy-related environments and might differ across cultures according to the type of orthographic script used by a child’s language (McBride-Chang et al., 2011). The phonological loop, on the other hand, is specialized for and actively involved in the processing and retention of verbal information, in the form of speech sounds or phonemes (Alloway, Gathercole, Willis, & Adams, 2004; Baddeley, Gathercole, & Papagno, 1998) and is related to phonological awareness skills and vocabulary development (Alloway et al., 2004; Baddeley, 1996, 2003; Baddeley et al., 1998; Hecht, Torgesen, Wagner, & Rashotte, 2001).

In addition to being related to phonological awareness and vocabulary development, the ability to rapidly process speech sounds, via the phonological loop, has also been linked to early literacy skills (Baddeley et al., 1998; Duyck, Szmalec, Kemps, & Vandierendonck, 2003; Gupta & MacWhinney, 1997; Jongejan, Verhoeven, & Siegel, 2007). Although the exploration of the visuo-spatial sketchpad has gained some attention in literature (Baddeley, 1996, 2003; Baddeley et al., 1998), the primary goal of the present study was to examine the relationship between EF and language development. As a result, when creating our task, we were less interested in accessing the visuo-spatial sketchpad and more focused on accessing the phonological loop. Considering that both EF and phonological awareness have been implicated in the development of children’s early literacy (Lan et al., 2011) and that the phonological loop has been specifically designated as the “language learning device” (Baddeley et al., 1998; Gupta & MacWhinney, 1997), we believed that focusing on the phonological loop was more relevant to our endeavor to examine the relationship between language development and EF.

Bearing in mind that (1) children’s executive processes develop at different rates (Anderson, 2002), (2) language development and EF impact one another, and (3) this relationship has clear implications for early literacy and academic achievement (Blair & Razza, 2007; Espy et al., 2006), it is *especially* relevant to advance a task that would

allow for examination of the relationship between language and EF, particularly in school-aged children. In addition, there continues to be enormous variability in the types of tasks used to assess EF skills (Carlson, 2005; Carlson & Meltzoff, 2008; Espy et al., 2006; Willoughby et al., 2010; Zelazo, 2006) and the components of EF that they assess. Although age-appropriate EF measures have been developed, assessing the rapidly developing cognitive skills of children has proved to be quite challenging (Anderson, 2002). For instance, there is little agreement as to which measures are most appropriate to use at any given age (Carlson, 2005). Often, tasks that are developmentally appropriate for one age group are not appropriate for another age group. As a consequence, recent efforts have been made to design EF tasks that are developmentally appropriate for children of different age ranges with varying levels of cognitive and language abilities (Carlson & Meltzoff, 2008; Espy et al., 2006; Willoughby et al., 2010). Some of the limitations of the most widely used EF tasks with children are that they (1) are not often sensitive enough to detect individual differences, (2) often lack the ability to differentiate specific cognitive deficits, (3) must be administered by highly trained research assistants in laboratory settings, and (4) are used with small, self-selected samples of children whose parents are motivated and willing to participate in university-based laboratory studies (Willoughby et al., 2010). Furthermore, many of the tasks used to test children's EF skills rely on dichotomous outcome variables (e.g., yes/no, pass/fail, right/wrong). Arguably, these "all or nothing" types of measures cannot accurately tap into nuances or individual differences in abilities. In addition, using small, self-selected samples makes it more difficult to generalize the results of these tasks to other types of populations (e.g., developmentally delayed, low-socioeconomic-status samples; Willoughby et al., 2010). This is precisely the reason that Blair, Zelazo, and Greenberg (2005) have asserted that careful psychometric research is needed using tasks that expose not only developmental trajectories, but also individual differences.

Due to these limitations, we had two goals for the present study: (1) to create a computerized EF task that would elucidate the relationship between language and EF, and (2) to validate this task by comparing the pattern of results to that of a well-established and widely used EF task (i.e., the Dimensional Change Card Sort task; Zelazo, 2006). The first goal of the present study was to create a task that would allow for the examination of the relationship between language and EF. For convenience, we computerized our EF task so that it could be administered with portable equipment at various locations and would not require administrators to have extensive training or expertise. This task, the lexical Stroop sort (LSS) picture-word task, allows for both dichotomous (e.g., correct/incorrect, pass/fail) and continuous measures (i.e., reaction time) to be recorded and

analyzed. Furthermore, the types of correct and incorrect responses can be analyzed for a more detailed examination of participants' patterns of responding. We validated this task using data collected from a diverse lower- and middle-class sample of 93 children enrolled in a public elementary school.

The second goal of the present project was to compare the patterns of results of the newly created LSS task to that of the well-established Dimensional Change Card Sort task (DCCS; Zelazo, 2006) to evaluate whether the LSS task is capable of providing insights into the relationship between language development and EF in children. The DCCS is an executive-functioning task that is suitable for use with children and has been used with both typically and atypically developing populations (Müller et al., 2008; Zelazo, 2006; Zelazo, Craik, & Booth, 2004; Zelazo, Müller, Frye, & Marcovitch, 2003). Several age-appropriate versions of the task exist: The standard version is typically used with preschool-aged children (i.e., 2.5 to 5 years of age) and requires participants to sort cards, by hand, first by one dimension (e.g., shape), and subsequently by another dimension (e.g., color). A more challenging extension of the task, the border game, is typically used with older, school-aged children (i.e., up to 7 years). In this more challenging version, children are required to sort by color or shape on the basis of the presence or absence of a border. If there is a border around the card, the children are asked to sort by one dimension (e.g., shape), but if there is not a border, they are to sort by the other dimension (i.e., color). See Fig. 1 for a visual depiction of stimuli for the DCCS task.

We created the LSS task to assess children's ability to sort information based on a particular dimension, similar to the DCCS task. Because our goal was to assess the cognitive processes that allow children to access their lexical and semantic knowledge bases, in developing our task, we took into account Baddeley's (1986) model of working memory, which incorporates the central executive, the visuo-spatial sketchpad, and most especially, the phonological loop. We were particularly interested in accessing the phonological loop in children's working memory because past research has indicated that phonological awareness is an important predictor of academic achievement (Blair & Razza, 2007; Jongejan et al., 2007). Also similar to the DCCS, the LSS task requires children to pay attention to a specific dimension of a stimulus and to ignore the irrelevant dimension to demonstrate their cognitive flexibility about rule-based dimensions of stimuli. However, unlike the DCCS, the LSS requires children to process the relevant and irrelevant information through their phonological loop and semantic knowledge base. Thus, we designed this lexical sorting task to specifically tap into children's language-processing skills, as well as their inhibitory control and cognitive flexibility.

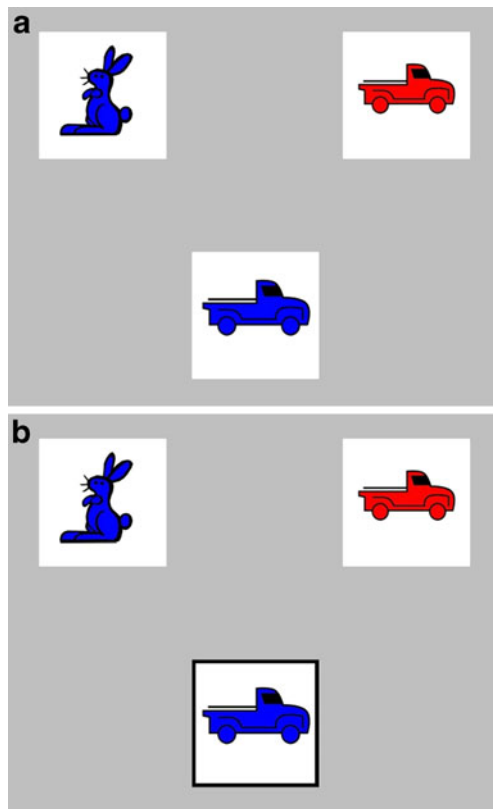
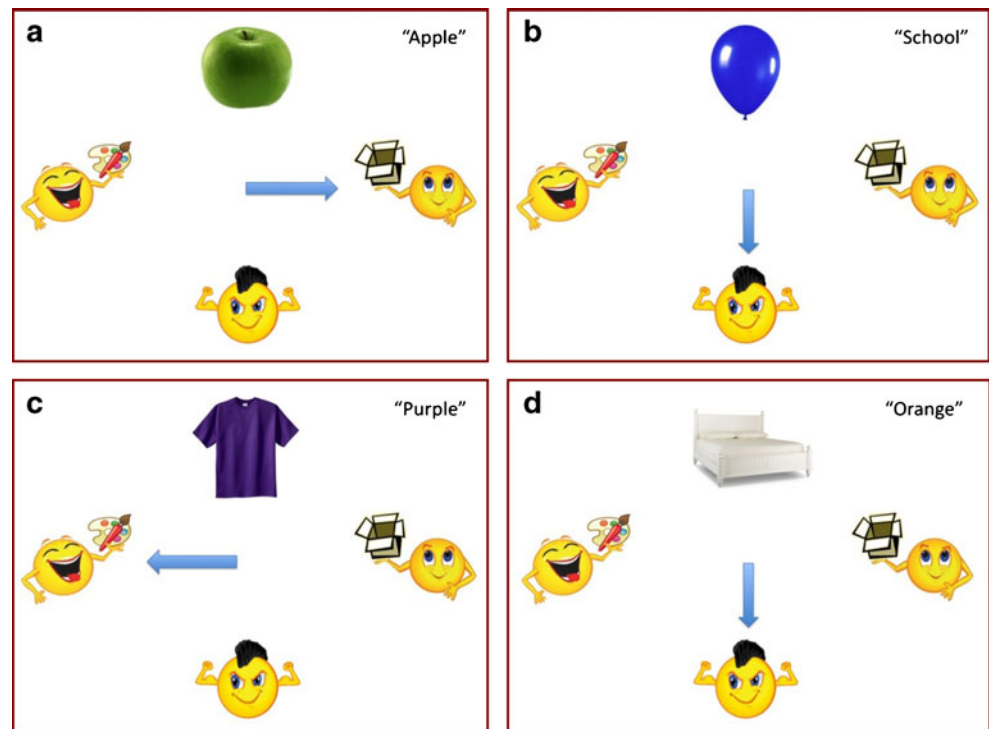


Fig. 1 Depiction of materials for the Dimensional Change Card Sort (DCCS) shape/color sort task (a) and border task (b). Adapted from Zelazo (2006)

The LSS requires participants to rapidly make decisions based on whether an auditory label does or does not match

Fig. 2 Depictions of the four possible sorting options for the lexical Stroop sort (LSS) task: (a) object match, (b) object mismatch, (c) color match, and (d) color mismatch. Words enclosed within quotation marks denote the presented auditory labels and were not visually displayed on the screen. Arrows denote the correct sorting selection for each example



a presented object. Then, participants must sort the object into the correct or target category. Specifically, in the LSS task, a conflict is created by a mismatch between the verbal label and the physical representation of that label in the image. For instance, children are simultaneously presented with a picture of an object (e.g., a green apple) and an auditory word (e.g., “red”) and must rapidly decide whether or not the picture (i.e., the physical representation of the word) matches the auditory word in terms of its label or color, by touching the correct portion of the screen. Thus, the task forces children to focus on the relevant dimensional relation (e.g., the physically represented object does not match the label) and to ignore/inhibit the conflicting dimensional relation (e.g., the color of the object does not match the word), all while keeping in mind the correct spatio-dimensional relationship in mind (place the image in the color gobbler vs. object gobbler vs. the mismatch gobbler). Considering that previous research with children has demonstrated that they are biased toward processing attribute similarity (e.g., the color and shape of objects) over spatial relations (Honomichl & Chen, 2011), the LSS task would create conflict by forcing them to consider both dimensional relations at the same time in order to respond accurately. See Fig. 2 for a visual depiction of the task.

In the present study, participants completed the LSS, standardized vocabulary, and phonological awareness assessments, as well as the DCCS. Since we sought to develop a task that specifically tapped into the relationship between language and EF, we were particularly interested in assessing the speed and efficiency of children’s lexical

access on the LSS. Previous research has shown that the speed with which children can access their lexical knowledge base is related to their developing phonological awareness and early literacy skills (Jongejan et al., 2007). Thus, based on the design and flexibility of the LSS task, we generated several hypotheses about children's performance on this task relative to their performance on the DCCS task in relation to their developing language abilities.

Our first hypothesis was that the relationship between reaction time and accuracy on the LSS would look different than the relationship between reaction time and accuracy on the DCCS. This hypothesis was based on the contrasting views presented in the literature about the roles of stimulus familiarity (Honomicl & Chen, 2011) and the congruency of stimulus characteristics (e.g., switching *red* color to *red* truck vs. switching from *red* color to *blue* truck; Kirkham et al., 2003) on processing speed and accuracy on switching tasks. Additionally, because the DCCS is largely a visual task and the LSS is an auditory task (i.e., differing modalities are exploited), it is difficult to make specific predictions regarding the relationship between reaction time and accuracy on the DCCS relative to the LSS. As a result, we were open to the possibility that children's reaction times on the DCCS might or might not be significantly related to their overall accuracy. However, for the LSS, we hypothesized that the children's accuracy would be related to their efficiency of lexical access. There were several possibilities we considered with regard to the direction of this relationship. The first was that children would generate slower reaction times, yielding higher accuracy (i.e., more processing time to make correct selections). In contrast, the second possibility was that children would generate faster reaction times, yielding lower accuracy (e.g., haphazardly responding). A third possibility was that children with faster reaction times would display higher accuracy, since they would be more efficient in accessing their semantic knowledge base, which would in turn facilitate their ability to make correct selections.

Our second hypothesis was that children's developing vocabulary skills would be indicative of their performance on both of the EF measures, thus demonstrating convergent validity. Previous research has shown that children's vocabulary development is related to the development of EF (Blair & Razza, 2007; McClelland et al., 2007; Singer & Bashir, 1999). As a consequence, in the present study, we predicted that children's receptive vocabulary skills would be positively related to their performance on both the LSS and DCCS tasks. More specifically, we predicted that children with larger receptive vocabularies would perform better on both tasks.

Our third hypothesis was that the children's phonological awareness skills would be related to their performance on the LSS, but not on the DCCS, thus demonstrating

discriminant validity. This prediction emerges from the theoretical impetus of the design of the LSS, in that it incorporates a substantial auditory or language-processing component (i.e., lexical access through the phonological loop). Unlike the DCCS, which is largely visual, the LSS has an overtly linguistic component (i.e., the verbal labels that require children to access their semantic knowledge base in specific categories: colors and objects) as well as auditory stimuli (i.e., the labels, distractors, and instructions are presented verbally) that the child must process simultaneously while trying to adhere to the rules associated with a specific gobbler (i.e., color gobbler vs. object gobbler vs. mismatch gobbler). Since so many verbal stimuli would be considered a heavy cognitive load for young school children's working memory capacity, we predicted that children's phonological awareness skills would be associated with their performance on the LSS. Also, because the DCCS does have a verbal component (i.e., the instructions are verbally provided) but is not as heavily dependent on verbal working memory, we predicted that children's performance on the DCCS would not correlate with their phonological awareness skills.

Method

Participants

The participants were recruited at a local elementary school servicing both middle and lower income communities in a southeastern city with approximately 64% of the student population receiving free or reduced price lunches. The total of 93 children (52 males, 41 females) between the ages of 5 and 8 years ($M_{\text{age}} = 6.95$ years, $SD_{\text{age}} = 0.89$ years, range = 4.9 to 8.4 years) participated in the study. The children were from various racial and ethnic groups (i.e., 34% African American, 29% White or Caucasian, 22% Hispanic, and 15% Other). Parental consent was obtained for all participants, and each child received stickers and pencils in appreciation for participation.

Of the 93 children who participated in the study, 4 (3 kindergartners, 1 second grader) were not included in the analyses of performance on the LSS task because they failed to meet the practice-trial criterion (described below). Thus, the final sample for these analyses was 89 children (50 males, 39 females).

For the DCCS, of the 93 children who participated in the study, 12 (9 kindergartners, 1 first graders, and 2 second graders) were not included in the analyses of performance on the DCCS task because they failed to reach the postswitch criterion, as outlined by Zelazo (2006) and described below. Thus, the final sample for these analyses was 81 children (47 males, 34 females).

For analyses comparing children's performances on the LSS relative to the DCCS, children who did not meet the inclusion criteria for both tasks were removed from the analyses ($n = 16$). As a result, the final sample for these analyses was 77 children (45 males, 32 females).

Apparatus and stimuli

E-Prime 2.0 computer software was used to create and administer the computerized versions of the LSS picture-word task and the DCCS task (Zelazo, 2006). Both tasks were administered using a PC laptop connected to a 15-in. ELO touch-screen computer monitor. The touch-screen monitor and E-Prime software were also used to record participants' responses (e.g., accuracy, reaction time). Auditory stimuli were presented through noise-reducing headphones (Sony Model MDR-7506).

LSS task Three sorting "gobblers" (i.e., yellow happy faces) were created using PowerPoint clip art and Adobe Illustrator. The "object gobbler" was depicted holding a box and was positioned midway down and on either the right or the left side of the computer screen. The "color gobbler" was depicted holding a color palette and a paintbrush and was positioned on the side of the screen opposite the object gobbler. The "mismatch gobbler" had a mischievous grin and a Mohawk hairstyle and was always positioned at the bottom center of the computer screen (see Fig. 2).

Demonstration stimuli Photographs of a blue fish, a green dress, a black car, and a purple flower were used during the demonstration phase. The blue fish was labeled "blue" (i.e., color match), the green dress was labeled "orange" (i.e., color mismatch), the black car was labeled "car" (i.e., object match), and the purple flower was labeled "school" (i.e., object mismatch).

Practice stimuli The labels used for the 16 practice trials are depicted in Appendix A.

Test stimuli Appendixes A and B present the test stimuli (labels and pictures) used in the present study. The test stimuli comprised 25 object labels, 20 object photographs, and 9 color labels. The 25 object labels were selected based on young children's familiarity with them. More specifically, we consulted the MacArthur Communicative Development Inventory: Words and Gestures (Fenson et al., 1993) to select object labels that referred to concrete nouns that children as young as 30 months were reported to both understand and say. We selected only labels that referred to concrete nouns to ensure that the corresponding objects could be clearly represented in still photographs (e.g., shoe, cup, book). Using this criterion, we generated a list of 25

words to use as object labels. Of these 25 labels, 5 were not paired with representative photos and always served as the object mismatch labels (i.e., "leg," "bag," "hat," "bottle," and "meat"). The remaining 20 object labels were paired with corresponding still photographs.

The 20 still photographs were selected because they were (1) easily recognized when presented in isolation and (2) monochromatic. Object photos were only selected if they could be clearly interpreted without additional cues. For example, the concrete nouns "chin" and "tongue" were difficult to visually represent without simultaneously showing the "mouth" attached to them. Similarly, the word "beach" was difficult to visually represent without simultaneously showing "sand" and "ocean." Furthermore, photographs were only selected if the depicted object could readily be identified as one color (e.g., *black* dog, *red* plane), so that clear color match and color mismatch labels could be appropriately applied to each photograph.

Naturally, the color match labels were selected based on the colors of their corresponding objects. As a result, nine different color labels were utilized. The color mismatch labels were selected and paired with objects based on two considerations. First, similarities in hue (e.g., purple/blue, red/orange, or black/brown) were considered, such that a red object would never be paired with the word "orange" as a color mismatch label. Secondly, similarities between the initial sounds of the words (e.g., "blue" vs. "black") were considered, such that a black object would never be paired with the word "blue" as a color mismatch label.

After selecting the 20 pictures and the corresponding object and color match/mismatch labels, we devised a counterbalancing strategy to ensure that each label category was equally represented across the 20 trials (i.e., 5 object matches, 5 object mismatches, 5 color matches, and 5 color mismatches). To accomplish this, we sorted the 20 pictures into four sets of 5 pictures while keeping three specific factors in consideration. The first factor was the conceptual categories of the objects. The concern was that if selection were left unconstrained, it was possible that several items from one conceptual category (e.g., animals) could be disproportionately paired with one type of label (e.g., object match labels). To minimize this possibility, objects in the same conceptual category were split across the four sets. For example, pictures of the dog, cow, and horse were assigned to separate sets, because they all fell under the category of animals. The specific categories taken into account when creating the list included animals, food, home/furnishings, and body parts/clothing.

The second factor considered in determining the sets was the objects' perceptual cues (i.e., color, shape), as represented in the pictures. Pictures of objects that were similar in shape (e.g., ball, cookie, apple) or color (light green apple vs. dark green cup) were assigned to

different sets. Again, the goal was to ensure that objects that were similar in size or color were not disproportionately paired with one type of label (e.g., color match labels). For example, two or more objects of a similar color (e.g., light blue vs. dark blue) were never presented with the same color match label (e.g., “blue”). Thus, for every set, the five objects selected were depicted in five different colors.

The third factor considered in determining the sets was the phonetic similarity of the labels. Specifically, similar-sounding words were not clustered within the same set. As a result, “cow,” “cup,” and “cookie” were intentionally separated due to the similarity in the initial velar /k/ sound. See Appendix A for a sample depiction of the counterbalancing strategy used in creating the five object sets, with corresponding object match, object mismatch, color match, and color mismatch labels.

DCCS task PowerPoint clip art pictures of red and blue rabbits and trucks were used as the stimuli for the DCCS. For the border version of the task, each combination (i.e., blue rabbit, red rabbit, blue truck, or red truck) had an identical counterpart surrounded by a 5-mm black border (Zelazo, 2006; see Fig. 2 for an example).

Peabody Picture Vocabulary Test—4th edition (PPVT-IV) The PPVT-IV is a standardized test of receptive vocabulary in English (Dunn & Dunn, 2007). Participants are presented with a sheet of paper or page that is broken up into four equal quadrants, each depicting an image. Participants are given an auditory label (e.g., “ball”) and instructed to select or point to the image that correctly corresponds to the label. Testing is concluded when a participant make eight incorrect responses within a 12-item section. The raw score is calculated by subtracting the number of errors committed in the entire assessment from the ceiling item (i.e., the last item). Per the PPVT-IV instruction manual (Dunn & Dunn, 2007), the raw score was calculated by subtracting the number of errors committed in the entire assessment from the ceiling item (i.e., the last item). Raw scores were translated into standardized scores according to the age of the participant. The PPVT is *standardized* with a mean score of 100 and a standard deviation of 15.

Comprehensive Test of Phonological Processing (CTOPP) The CTOPP is a standardized measure of phonological processing comprising several subtests (Wagner, Torgesen, & Rashotte, 1999). For the present study, we used the Phonological Awareness subtest. This subtest is composed of two tasks: elision and blending words. During the elision task, children heard a word, and then had to repeat it while omitting one of the sounds in the word. For example, when hearing the word “tiger,” participants had to repeat it without the “g” (i.e., \tī- r̄). During the blending-words task,

participants combined various sounds in order to form a word. Participants heard a word broken up into individual sounds (e.g., /t/ + /oi/) and had to merge the sounds to form the given word (i.e., “toy”). Raw scores are calculated by counting the number of correct responses for each subset (i.e., elision and blending). To calculate the standard scores, the children’s scores on the subsets were added together to create a total raw score of their phonological awareness. This score was then transformed into a standardized score, as indicated by the instructional manual (Wagner et al), based on the relevant children’s age group (i.e., 5–11 years). Similar to the PPVT, the composite score for the CTOPP is *standardized* with a mean score of 100 and a standard deviation of 15.

Procedure and coding

After receiving parental consent, children were tested individually during school hours. A female experimenter tested each participant in a small, quiet room at the school. Children were tested during two sessions on two different days. Before each session, participants were informed that they could stop the session and return to their classroom without any consequence at any time. After child assent was given, the testing session began. Each session lasted approximately 20 min.

For both the DCCS and LSS tasks, children were seated in front of the touch-screen monitor and alongside an experimenter who was positioned in front of the laptop. The DCCS and LSS tasks were administered on different days in order to reduce the possibility that performance on one task would directly influence performance on the subsequent task. As a result, children were given the DCCS and the LSS in two different sessions with at least a 1-week gap between. The order in which the tasks were given (e.g., LSS first, DCCS second, or vice versa) was counterbalanced across participants.

LSS task For the LSS task, children were seated in front of the touch-screen monitor, next to the experimenter. The experimenter initiated the task by depressing a key on the laptop connected to the monitor. First, children were introduced to the three “gobblers.” Children were told that the object gobbler was moving and needed help packing the correct objects in his box. The color gobbler was painting a picture, but was running low on paint and needed help collecting colors for her paint palette. The mismatch gobbler liked to trick people and was trying to prevent the other gobblers from being helped. Second, children were informed that it was their job to help the object and color gobblers while avoiding being tricked by the mismatch gobbler (see Appendix C for the instructions). Finally, children were told that they were going to see a picture and hear a word. If the word matched the

color of the picture, they were to touch the color gobbler. If the word matched the *object*, they were to touch the object gobbler. However, since the mismatch gobbler liked to switch the pictures around, sometimes the word would not match the object or the color, so they were to “give it back to him” by touching the mismatch gobbler.

Demonstration phase After the rules of the game were explained, the experimenter invited the participant to play the game with her for four demonstration trials (color match, object match, color mismatch, and object mismatch examples). During this phase, the experimenter narrated the task and asked the child for his/her help.

Practice phase After the demonstration phase, children were initially given 16 practice trials (4 from each label type category). The experimenter watched and recorded the numbers of correct and incorrect responses. If a child scored significantly above chance (i.e., got 10 out of 16 or 62.5% correct responses), he or she proceeded to the test phase. If a child got fewer than 10 correct, he or she was reminded of the instructions and allowed to complete 8 more practice trials that were identical to 8 of the 16 trials previously seen during practice. It should be noted that 70% (i.e., 65 of 93) of the participants met or exceeded this criterion. Only 24 of the 93 children completed the additional practice trials, over half of which were kindergartners ($n = 14$). Of these 24 children who had extra practice trials, 4 (3 kindergartners, 1 second grader) were removed from all analyses for the LSS because they failed to respond correctly at above chance levels (i.e., 5 out of 8 correct responses, or 62.5%).

Test phase After the practice phase, children proceeded to the test phase. The experimenter told participants that it was now their turn to play the game by themselves with the headphones on. Children were informed that the experimenter could not help them during the task because she was unable to hear the words. If they did not hear or know a word, they were instructed to make their best guess and keep going. Once children put on the headphones, a volume check trial was conducted.

During this volume check trial, a prerecorded audio file instructed participants to touch a specific gobbler (e.g., “touch the object gobbler”), which was also designated by an arrow. Once the participants had correctly touched each of the three gobblers, as instructed by the audio and arrows, the experimenter removed their headphones and asked if the volume was at an appropriate level. Next, the experimenter reminded participants about the rules of the game and informed them that their responses would be timed, so they should try to make their choices as quickly as possible. Once a participant acknowledged that he or she was ready to play the game, the test phase was administered.

During the test phase, participants completed 20 trials. The presentation order of the trials was randomly determined. While the mismatch gobbler was always on the bottom center of the screen, the left/right orientations of the object gobbler and the color gobbler were also counterbalanced across all participants (see Fig. 2). In order for a response to be recorded, the participant had to touch somewhere within a 3-in. radius of a gobbler. A response made outside of this radius was recorded as “no response.” A trial did not proceed unless a touch somewhere on the screen was registered. The entire task took approximately 5–7 min to complete.

Coding Children’s responses during the test phase were coded as correct, incorrect, or “no response.” Correct, incorrect, and no responses for each of the sorting categories (i.e., color match/mismatch, object match/mismatch) were tallied separately. These tallies were converted to percentage scores. The percentages of correct and incorrect responses in each sorting category were calculated by adding up the number of correct or incorrect responses and dividing by five, minus the number of “no responses.” Across all categories, the total numbers of correct responses, incorrect responses, and no responses for the match and mismatch trials were calculated. The percentage accuracy was calculated by dividing the number of correct responses over 20 trials, minus the total number of “no responses.” Reaction time, measured in milliseconds, from the initial presentation of the object and auditory label to response/selection on the touch screen, was recorded via the E-Prime computer software.

DCCS task For the DCCS, we followed the procedure outlined by Zelazo (2006), except that we created a computerized version using children’s fingerpresses on the touch-screen monitor as the behavioral response. Following Zelazo’s (2006) protocol, children were given two demonstration trials during which they were introduced to the stimulus pairs (e.g., “Here’s a red truck and here’s a blue rabbit”) and told about the sorting game (e.g., “This is the shape game”; see Zelazo, 2006, for detailed instructions). On the computer screen, the two pictures (i.e., rabbit and truck) were presented side by side at the top of the screen. At the bottom of the screen, a new “card” (i.e., a picture of a truck or rabbit) was presented (see Fig. 1). During this demonstration phase, the experimenter sorted the upcoming pictures to the appropriate place on the monitor while explaining these actions (e.g., “See, here’s a rabbit. So it goes here [touches rabbit on left side of monitor].”) After this demonstration phase, children proceeded to the preswitch phase.

During the preswitch phase, children were given six trials to sort by the predefined rule (e.g., shape game = rabbits on the right, trucks on the left). The presentation of the six trials was randomized. After completing the preswitch phase, children proceeded to the postswitch phase. During this phase, children

were instructed to sort again, but this time by a different rule (e.g., sort by color *instead* of shape). Again, the presentation of the six postswitch trials was randomized. After completing the six postswitch trials, children proceeded on to the border version of the task.

During the border version of the task, children were asked to sort the cards that had a border by color and the cards with no border by shape (see Fig. 1). For example, if the picture of the red rabbit had a border around it, children were to play the “color game.” However, if there was no border, children were to play the “shape game.” After these instructions were explained to the participants, the experimenter encouraged them to do their best and to always make a selection, even if they were unsure of the answer. At this point, the experimenter answered any questions, then instructed the participant to put on the headphones to begin the task. Before each trial, prerecorded verbal instructions reminding the participant how to sort the cards were repeated through the headphones (Zelazo, 2006). A trial ended only when the participant made a selection. The task was complete after the child completed 12 trials (see Zelazo, 2006, for a detailed description of the task). For the present study, the left/right orientations of the objects (rabbit/truck) on the monitor and the colors (red/blue) of each object were counterbalanced across all participants. The entire task took approximately 5–7 min to complete.

Coding For the preswitch phase, the numbers of correct and incorrect responses out of 6 trials were recorded. For the postswitch phase, the numbers of correct responses and perseverative errors (i.e., selections made based on the initial rule) out of six trials was recorded. Children who responded correctly on at least five trials were classified as “passing” the postswitch phase (Zelazo, 2006). Using this criterion, 12 children did not pass this phase and thus did not continue on to the border version of the task. For the border version of the task, the numbers of correct and incorrect responses out of 12 trials were recorded. Children who responded correctly on 9 out of the 12 (75%) trials were classified as “passing” the task (Zelazo, 2006). Also, for the DCCS, reaction time (in milliseconds) was coded from the presentation of the object or “card” to the instance of a registered touch/selection. Reaction time was recorded via E-Prime computer software.

Results

LSS task

Table 1 presents the means and standard deviations for children’s accuracy and reaction times on the LSS as well as their standardized scores on the PPVT and CTOPP, along with

Table 1 Partial correlations between accuracy and reaction times on the LSS and standardized PPVT and CTOPP scores, controlling for age, as well as the means and standard deviations (*SDs*) for all measures

Variables	1	2	3	4	Mean (<i>SD</i>)
1. LSS % correct	–				83.4% (20.8)
2. LSS reaction time	–.34***	–			2,562.3 ms (538.49)
3. PPVT	.53***	–.17	–		101.0 (16.80)
4. CTOPP	.32**	.11	.45***	–	105.0 (13.14)

** $p < .01$. *** $p \leq .001$

the partial correlations, controlling for children’s age, between all of these measures. A one-way analysis of covariance (ANCOVA) was conducted on children’s percentages correct on the LSS task, with gender as a between-subjects factor and age as a covariate. The findings failed to revealed a significant main effect of gender, $F(1, 86) = 0.71, p > .10$.

As predicted, the correlation analyses revealed that children’s reaction times were significantly related to accuracy, although the magnitude of the relationship was moderate in size ($r = -.34$). More specifically, faster reaction times were related to more correct responses, whereas slower reaction times were related to more incorrect responses. In other words, children who responded more rapidly were more likely to be accurate, and children who responded less rapidly were more likely to be inaccurate in their responses. To further assess the relationship between reaction time and accuracy, we conducted a logistic regression to determine whether children’s reaction times predicted whether they passed or failed the task at a 75% accuracy criterion (Zelazo, 2006). The findings revealed that children’s reaction times did reliably predict whether they passed or failed the task, $\beta = -.001, Wald(86) = 5.94, p < .05$: Children with faster reaction times were more likely to pass the task.

Since the goal of the study was to develop a measure that would allow for the examination of the relationship between EF and language development, we conducted additional analyses to determine the interrelations between children’s oral language scores and performance on the LSS. The partial correlations (see Table 1) revealed that children’s scores on the PPVT and CTOPP were significantly related to their accuracy on the LSS; children with higher PPVT and CTOPP scores were more likely to generate more accurate responses.

Based on the findings demonstrating that children’s PPVT scores were correlated with their CTOPP scores, and that these two variables were in turn associated with performance on the LSS task, we explored the presence of a mediating effect. Specifically, in line with our theoretical impetus for developing the LSS task, we tested whether children’s vocabulary mediated the relationship between their phonological awareness and their performance on the

Table 2 Multiple regression predicting children's performance on the LSS task

Variable	LSS		β
	<i>B</i>	<i>SE B</i>	
Constant	– 0.052	0.170	
Age	0.009	0.011	.078
CTOPP	0.002	0.002	.101
PPVT	0.006	0.001	.518***

$R^2 = .32$, *** $p < .001$

LSS task. Statistically, mediation is said to be present when it can be demonstrated that the impact of one variable on another is through a third variable, which is referred to as a *mediator* (Hoyle & Kenny, 1999).

Statistical evidence of a mediation effect requires confirmation that (a) the independent variable (phonological awareness) has a causal influence on the dependent variable (performance on the LSS task) and (b) the independent variable has a significant indirect effect on the dependent variable through the mediator (vocabulary), indicated by a decline in the effect of the independent variable on the dependent variable when the mediator is accounted for (Hoyle & Kenny, 1999). To test for mediation, we used the causal-steps approach (Preacher & Hayes, 2008) and ran a series of regression analyses. As recommended by Baron and Kenny (1986), we first regressed the mediator (vocabulary) on the independent variable (phonological awareness), then regressed the dependent variable (performance on the LSS task) on phonological awareness. Then we regressed performance on the LSS task (i.e., percent correct) on both the children's vocabulary and their phonological awareness. Children's age was entered in the first step of all regression models as a control variable.

The first regression model was significant ($R^2 = .22$), $F(2, 86) = 12.16$, $p < .001$, indicating that the independent variable (phonological awareness) predicted the mediator (vocabulary). The second regression analysis was also significant ($R^2 = .13$), $F(2, 86) = 6.57$, $p < .01$, demonstrating that phonological awareness predicted performance on the LSS task. The final regression model (see Table 2) was also

significant ($R^2 = .34$), $F(2, 85) = 14.69$, $p < .001$, and children's PPVT scores, $t(86) = 5.19$, $p < .001$, uniquely predicted their performance on the LSS. Furthermore, the effect of phonological awareness was significantly reduced from the second ($\beta = .34$) to the final ($\beta = .10$) regression model (R^2 change = .21; $F = 26.96$, $p < .001$).

In order to test the significance of the mediation model, we used the online statistical program MedGraph (Jose, 2003), which indicated that a full mediation effect was present, with a significant Sobel z value of 3.80 ($p < .001$). The total effect (i.e., the sum of both the direct and indirect effects of phonological awareness on children's performance on the LSS task) revealed that 34% of the variance in the children's LSS scores was explained by the model. Although correlation coefficients and standardized regression coefficients often serve as measures of effect sizes of paths in a mediation model, the ratio of the indirect to the total effect is also used to provide insight (MacKinnon, Fairchild, & Fritz, 2007). An examination of the ratio of indirect to total effect indicated that 71% of the variance in children's LSS scores explained by the model was accounted for by the mediating path from children's phonological awareness through their vocabulary. An additional 29% of the variance was accounted for by the direct path between phonological awareness and the children's performance on the LSS (see Fig. 3). The findings of these analyses indicate that our predictions about the LSS task were supported, as children's performance on the task was dependent on both their vocabulary and their phonological awareness skills. Since the LSS task incorporates a substantial auditory or language-processing component, children had to use their phonological loop to access their lexical base.

DCCS

Only children ($n = 81$) who passed the postswitch phase of the task (Zelazo, 2006) were included in the following analyses. A preliminary analysis revealed no effect of sorting dimension order (i.e., color or shape first) on any of the outcome measures for the preswitch, postswitch, or border versions of the task, all $ps > .10$. Thus, data were collapsed

Fig. 3 Depictions of the mediating effect of children's vocabulary on the relationship between phonological awareness and accuracy in the LSS task; direct effect = .24, indirect effect = .10. ** $p < .01$. *** $p < .001$

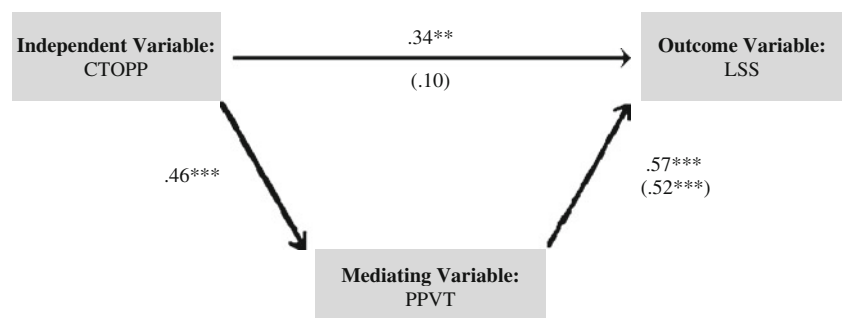


Table 3 Partial correlations between accuracy and reaction times on the DCCS and standardized PPVT and CTOPP scores, controlling for age, as well as the means and standard deviations (*SDs*) for all measures

Variables	1	2	3	4	Mean (<i>SD</i>)
1. DCCS % correct	–				84.6% (11.2)
2. DCCS reaction time	.04	–			3,589 ms (786.7)
3. PPVT	.37**	.03	–		102.8 (13.21)
4. CTOPP	.20	.23*	.39***	–	105.6 (13.21)

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

across this variable. The first set of analyses sought to determine whether the present sample yielded the anticipated patterns of results on the DCCS (Zelazo, 2006). Zelazo reported that the majority of 5-year-olds pass the postswitch phase, with at least 5 out of 6 correct responses. Data from the kindergartners (i.e., 5- to 6-year-olds) in the present study revealed that the majority (i.e., 28 out of 37, or 76%) passed the postswitch phase. Of the 5-year-olds who typically pass the postswitch phase, Zelazo reported that about 50% go on to pass (i.e., at least 9 out of 12 correct, or 75%) the border version of the task. In the present study, as predicted, about 54% (i.e., 15 out of 28) of the kindergartners who passed the postswitch phase also passed the border version of the task. Taken together, the patterns of results for the present study are congruent with patterns traditionally found with younger children on the DCCS (Zelazo, 2006).

Table 3 presents the means and standard deviations for children's accuracy and reaction times on the DCCS as well as their standardized scores on the PPVT and CTOPP, along with the partial correlations, controlling for children's age, between all of these measures. A one-way ANCOVA was conducted on children's percentages correct on the DCCS task, with gender as a between-subjects factor and age as a covariate. A main effect of gender, $F(1, 78) = 6.59$, $p < .05$, $\eta_p^2 = .08$, demonstrating that the girls ($M = 87.6\%$, $SD = 11.9\%$) generated significantly more correct responses than did the boys ($M = 82.5\%$, $SD = 10.2\%$). However, both boys and girls generated significantly more correct than incorrect responses (all $ps < .05$).

The correlation analyses revealed that children's reaction times were not significantly related to accuracy on the DCCS. The findings also revealed significant relationships between children's accuracy and their scores on the PPVT. Children's PPVT scores were positively related to their accuracy on the DCCS (i.e., higher vocabulary scores being associated with more correct responses). Children's scores on the CTOPP were not related to their accuracy on the DCCS, with correlations approaching marginal significance ($p = .07$). We conducted a linear regression analysis to examine whether PPVT scores predicted children's performance on the DCCS. Children's age was entered into the

regression equation in the first step as a control variable. The independent variable was children's PPVT scores, with the dependent variable being their percentages correct on the DCCS. The findings revealed that the overall regression model was significant, $R^2 = .14$, $F(3, 77) = 4.10$, $p < .01$, with children's PPVT scores uniquely predicting their performance on the DCCS, $\beta = .56$, $p < .001$. No association was found between the children's performance on the DCCS and their age, $\beta = -.02$, $p > .10$.

Validating the LSS

Another goal of the study was to compare the results of the LSS and the DCCS. To do this, we conducted an additional set of analyses comparing children's performance on the LSS relative to their performance on the DCCS. Because the numbers of children who successfully completed the two tasks differed, it was important to ensure that we compared scores only for children who both passed the postswitch phase on the DCCS and met the criterion during practice trials for the LSS; only children who met the inclusion for both tasks were included in the analyses ($n = 77$). Of these 77 children, 53 (i.e., 68.8%) passed both tasks (see Table 4).

To further explore the degree to which children's performance on the LSS yielded patterns of responding similar to those on the DCCS, we conducted a series of analyses. Preliminary analyses using gender as a between-subjects variable failed to yield significant results, all $ps > .10$, and as a result, all subsequent analyses were collapsed across gender. A related-samples McNemar test revealed that the proportion of children who passed the LSS with 75% or more correct responses did not significantly differ from the proportion of children who passed the DCCS task, $n = 77$, $p = .66$. In addition, the findings revealed that children's percentages correct on the LSS ($M = 86.8\%$, $SD = 18.0\%$) did not significantly differ from their percentages correct on the DCCS ($M = 84.6\%$, $SD = 11.4\%$), $t(76) = -1.05$, $p > .10$. Thus, children displayed similar patterns of responding for the LSS and the DCCS, with almost equal percentages of correct responses.

Furthermore, we conducted correlation analyses and found a positive and significant relation between children's accuracy

Table 4 Frequency distribution of student performance on the LSS and DCCS tasks

LSS	DCCS		Total
	Pass	Fail	
Pass	53	12	65
Fail	10	2	12
Total	63	14	77

(percent correct) on the LSS and their accuracy (percent correct) on the DCCS, $r = .25$, $p < .05$. More specifically, children with higher levels of accuracy on the LSS also had higher levels of accuracy on the DCCS, demonstrating some shared variance between the two measures.

A 2 (gender) by 2 (reaction time: LSS vs. DCCS) mixed-model ANCOVA controlling for age revealed that children took significantly longer to respond on the DCCS ($M = 3,585.44$ ms, $SD = 783.12$) than on the LSS ($M = 2,523.80$ ms, $SD = 533.30$), $F(1, 74) = 4.67$, $p < .05$, $\eta_p^2 = .06$. No other main effects or interactions emerged significant, all $ps > .10$. Partial correlations, controlling for children's age, revealed a significant and positive association between children's reaction times on the LSS and the DCCS ($r = .50$, $p < .001$).

Discussion

While there continues to be widespread interest in executive function, due to its implications in brain development (Miller & Cohen, 2001), cognitive and socioemotional development (Carlson & Moses, 2001; Hughes & Ensor, 2007), school readiness (Blair, 2002), and academic achievement (Blair & Razza, 2007; Gathercole, Pickering, Knight, & Stegmann, 2004), not much is known about the relationship between EF and language development in school-aged children. The purpose of the present project was to create and validate a computerized EF task that would allow for a direct examination of the relationship between language and EF. This is especially relevant because both of these cognitive skills are implicated in academic achievement for use with school-aged children. In order to attain both convergent and discriminant validity for the measure, we tested a racially and economically diverse sample of children 5–8 years of age with both the LSS task and the widely used DCCS task (Zelazo, 2006). We used additional standardized measures to test the children's receptive vocabulary and phonological awareness skills in order to assess the associations between children's EF and oral language skills.

Our first prediction was that children's reaction times on the LSS task would be related to their accuracy on the task. Because of the manner in which the stimuli in the LSS task were presented (i.e., orally) and the type of stimuli (i.e., words commonly known by children), we speculated that the LSS would require children to quickly access their semantic knowledge base through the phonological loop in working memory. Thus, we assumed that children who had better phonological awareness skills would respond faster on the task, particularly since vocabulary and phonological awareness are related to one another (e.g., Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003). As a result, it was possible that children who were

faster would also be more accurate. So, we had assumed that the reaction time data for the LSS task would be able to provide information about the speed with which children access their semantic knowledge base, and the children's accuracy on the task would provide information about the efficiency with which they were able to access their semantic knowledge base. The data provided support for this claim. Specifically, children's reaction times were related to their accuracy on the task and also predicted whether they were likely to pass or fail the task. Children who responded faster during the task were (1) more likely to be more accurate and (2) more likely to "pass" the task with 75% or more correct responses. Thus, assessing reaction time and accuracy measures on the LSS task allowed for a more comprehensive view of children's developing cognitive abilities, indicating that the task does indeed provide information about both the speed and efficiency of children's cognitive and oral language skills.

Considering that our task was designed to provide insight into the relationship between EF and language development in school-aged children, we also predicted that children's scores on the LSS would be related to their scores on the standardized oral language measures, the PPVT and the CTOPP. Given the constraint that the LSS task incorporates a substantial verbal component (i.e., phonological processing in lexical access), we were especially interested in the relationship between children's scores on the LSS task and their scores on the standardized task assessing phonological awareness. Our prediction was that children's phonological processing skills would be utilized in the LSS task, which would be demonstrated by a unique association between the children's accuracy on the LSS and their phonological awareness skills. The findings of the mediation analysis confirmed this prediction; phonological awareness had both direct and indirect (through vocabulary) causal impacts on the children's performance on the LSS, explaining 32% of the variance in their scores. Our study provides additional support for previous studies that have found an association between preschool and kindergarten children's phonological awareness and EF (e.g., Blair & Razza, 2007). But we were also able to add to the current literature in two ways: First, we extended the relationship between phonological awareness and EF to older school-aged children, and second, we developed a causal model that elucidated the relationship between phonological awareness and EF.

Another goal of the study was to validate the LSS with a widely used task of EF, the DCCS. Toward this end, we predicted that we would find a similar pattern of relationships between children's vocabulary and the children's results on the LSS and the DCCS tasks. As expected, both tasks were similarly related to children's scores on the PPVT. This finding replicates those from previous studies that have demonstrated that children's receptive vocabulary

(as measured by the PPVT) is associated with their EF skills (e.g., Blair & Razza, 2007). Most importantly, children's scores on the LSS were correlated with their scores on the DCCS. Together, these findings provide convergent validity for the LSS task in relation to the DCCS task and indicate that the LSS task does indeed tap into cognitive abilities similar to those tapped by the DCCS task.

However, we did not find a significant relationship between the children's scores on the DCCS task and their phonological awareness. Interestingly, children's reaction times on the DCCS were positively correlated with their phonological awareness skills, indicating that children who had better phonological awareness skills took longer to respond on the DCCS task. Research has shown that children with higher phonological awareness are likely to have improved representations of language (Dickinson et al., 2003) and, as a consequence, may be better at representing complex rule structures (Blair & Razza, 2007). Since the DCCS task requires children to direct their attention resources to relational rules that determine how stimuli are to be sorted, it is entirely possible that children with better phonological awareness tried harder to overcome their attentional inertia (Honomichl & Chen, 2011), which resulted in longer processing times on the task. However, since children's reaction times were unrelated to their accuracy on the DCCS, we cannot be certain whether the children who took longer were also the ones who performed better (or worse) on the task.

Previous research has shown that stimulus attributes can bias attentional resources in switching tasks (Honomichl & Chen, 2011), and since the DCCS requires children to allocate attentional resources to representing complex rule structures, taking longer on the task may actually have increased bias rather than minimized it. This might be the reason why we did not find a significant association between reaction times and accuracy on the DCCS. Overall, the results indicate that although the DCCS and LSS tasks access similar cognitive abilities in children, they do not tap into identical cognitive processes. The differing patterns of relationships between children's phonological awareness and the two tasks provide discriminant validity for the LSS.

Both the LSS and DCCS tasks assess children's ability to pay attention flexibly to rule-based dimensions of objects. In order to succeed in the LSS task, children had to accurately categorize the stimuli according to attribute similarity (i.e., does the color or object match the auditory label?) and relational similarity (i.e., does it go into the color match gobble, the object match gobble, or the mismatch gobble?) simultaneously. But cognitive processes that children use to succeed in the LSS task differ from those used in the DCCS task. Unlike the DCCS, the LSS task was

specifically designed to access the relationship between children's language development and executive function. As such, the mediation model provides evidence to suggest that both phonological awareness and vocabulary influenced the children's performance on the LSS task. The children processed the relevant and irrelevant information related to the dimensions of the objects they were required to sort by accessing their vocabulary through their phonological loop. As Fig. 3 depicts, stimuli in the LSS task activated the children's phonological loop. This activation of the phonological loop facilitated children's access to their vocabulary. This access directly influenced children's ability to use their vocabulary to make selections on the LSS task. These findings provide evidence suggesting that, in addition to assessing children's cognitive flexibility, the LSS is able to tap into the relationship between language development and EF above and beyond the cognitive processes tapped into by the DCCS.

Finally, another goal of this project was to address some of the limitations put forth in the current literature pertaining to the problems associated with using EF tasks with children at varying developmental stages (e.g., Anderson, 2002; Willoughby et al., 2010). To address this problem, we created a task that tapped into EF in children across a broad range of ages and was still be sensitive enough to detect individual differences, particularly in regard to the relationship between language and EF. Furthermore, research assistants with little training can administer this task in various settings. In addition to tapping into children's phonological awareness skills, the LSS also allows for exploration of individual differences through the use of both categorical and continuous measures.

While the validation of this task was conducted with school-aged English-speaking children, the utility of the LSS allows for testing populations with different language environments and developmental patterns. For example, the auditory labels could be substituted with age-appropriate vocabulary in multiple languages, thereby extending the use of this task to bilingual populations. Thus, the LSS could easily be adapted with culturally relevant stimuli, thereby providing an opportunity to explore differences between individuals of varying language environments (e.g., native Spanish speakers vs. instructional Spanish speakers, or monolinguals vs. bilinguals; Bialystok, 2001). Furthermore, this task could be used to explore the relationship between numerical cognition, number words, and EF by substituting number words and numbers for the words and objects (Hecht et al., 2001). Although the primary goal of creating this task was to explicitly tap into the phonological loop, the design of the task allows for a more direct examination of the role of the visuo-spatial sketchpad. More specifically, the development of children's visuo-spatial sketchpads

could be studied by randomly changing the spatial locations of the correct and incorrect “gobbler” choices in relation to the auditory stimuli. Future research will be needed in order to arrive at any firm conclusion about the efficacy of this task in providing such information.

In addition, the LSS may provide insight into both typically and atypically developing populations. By replacing the object labels and pictures used in the present study with letter words (e.g., “A”) and letters (e.g., upper and lowercase), the task may be utilized to explore early phonological or letter knowledge in preliterate preschoolers and kindergartners. Likewise, written text (e.g., the word “apple”) could be substituted for the pictures of objects to explore early orthographic knowledge and developing literacy skills. Furthermore, research has shown that phonological processing speed is a more accurate predictor for reading skills and disabilities than phonological awareness (Jongejan et al., 2007; Wagner et al., 1997). Therefore, the processing speed and accuracy measures on the LSS might provide early indicators of children who are at risk for developing poor reading skills or dyslexia. Given the greater ease and

efficiency of administering the LSS relative to traditional phonological awareness tests, such as the CTOPP, the LSS may prove to be a beneficial tool for screening in the initial stages of identification. Thus, a task for future research is to assess the nature of the relationship between performance on the LSS (i.e., speed and accuracy) and early literacy and developing reading abilities. The utility and adaptability of the LSS presents a convenient and novel approach for assessing executive-functioning abilities in various populations, unparalleled by current commonly used measures in the field.

Author note Equal contributions were made by L.E.K. and V.K. This research was supported by a Language Learning Research Grant and a Ralph E. Powe Junior Faculty Enhancement Award awarded to M.P.W. in 2010. We thank Maria Paula Daneri, Maria Laura Valdes, Jenny Wang, and Kathleen Ghio for their assistance in data collection, coding, and data entry. We also sincerely thank Miles Plumlee for his assistance in participant recruitment. In addition, we are greatly indebted to Ari Cohen as well as to the faculty and staff at SWE. We also thank Keith Whitfield, Laura Kuhn, and the WILD writing group for all of their helpful suggestions at various stages of creating the measure, as well as on previous drafts of the manuscript. Finally, we express our sincerest gratitude to all of the students and families who participated in the study.

Appendix A

Lexical Stroop Sort auditory–visual stimuli

Table 5 Practice trials: Pictures and labels

Pictures	Label Presented	Sorting Category
Black Car	“Black”	Color Match
Green Dress	“Green”	Color Match
Blue Plate	“Blue”	Color Match
White Duck	“White”	Color Match
Purple Flower	“Orange”	Color Mismatch
Brown Bread	“Green”	Color Mismatch
Black Bear	“Red”	Color Mismatch
Yellow Pencil	“Blue”	Color Mismatch
Tan Ear	“Ear”	Object Match
Tan Foot	“Foot”	Object Match
Brown Ice Cream	“Ice Cream”	Object Match
Red Spoon	“Spoon”	Object Match
Tan Hand	“School”	Object Mismatch
Tan Nose	“Paper”	Object Mismatch
Orange Crayon	“Flag”	Object Mismatch
Blue Fish	“Cake”	Object Mismatch

Boldface text indicates the 8 trials that were repeated if a child scored less than 11 out of 16 trials correct.

Table 6 Test trials: Pictures and labels

Pictures	Object Labels	Color Labels
Black Cow	“Cow”	“Black”
Blue Balloon	“Balloon”	“Blue”
Orange Shoe	“Shoe”	“Orange”
Purple House	“House”	“Purple”
Yellow Cheese	“Cheese”	“Yellow”
Brown Cookie	“Cookie”	“Brown”
Green Table	“Table”	“Green”
Red Airplane	“Airplane”	“Red”
White Tooth	“Tooth”	“White”
Black Dog	“Dog”	
Orange Ball	“Ball”	
Brown Horse	“Horse”	
Green Cup	“Cup”	
Red Phone	“Phone”	
Yellow Chair	“Chair”	
Blue Book	“Book”	
Green Apple	“Apple”	
Purple Shirt	“Shirt”	
Red Mouth	“Mouth”	
White Bed	“Bed”	
	“Hat” *	
	“Leg” *	
	“Bottle” *	
	“Meat” *	
	“Bag” *	

*Only used with the object mismatch sorting parameter.

Table 7 Example of picture–label sorting assignment counterbalancing strategy

Pictures	Corresponding Labels
Color Match	
Black Car	“Black”
Green Dress	“Green”
Blue Plate	“Blue”
White Duck	“White”
Color Mismatch	
Purple Flower	“Orange”
Brown Bread	“Green”
Black Bear	“Red”
Yellow Pencil	“Blue”
Object Match	
Tan Ear	“Ear”
Tan Foot	“Foot”
Brown Ice Cream	“Ice Cream”
Red Spoon	“Spoon”
Object Mismatch	
Tan Hand	“School”
Tan Nose	“Paper”
Orange Crayon	“Flag”
Blue Fish	“Cake”

Appendix B

Select photograph examples of the test phase visual stimuli: objects and colors



Appendix C

Instructions for the Lexical Stroop Sort task

Let’s play a game. These are some of my friends, the gobblers.

This is the color gobbler. She wants to paint a picture but is running low on paint. She needs your help with collecting colors for her paint palette.

This is the object gobbler. He is moving to another town and is busy packing up his things in a cardboard box. He needs your help making sure he gets all of the right objects in his box so he doesn’t forget anything.

This is the mismatch gobbler. He likes to trick people. He has heard that the other two gobblers have asked for your help and he wants to get in the way!

It is your job to help these two gobblers without being tricked by the mismatch gobbler.

During the game, you are going to hear a word and see a picture.

If the word matches the COLOR of the picture, touch the color gobbler. So if you hear the word “red” and there is a picture of a red motorcycle, touch the color gobbler as quickly as you can to give her that color for her paint palette.

If the word you hear matches the name of the object, touch the object gobbler. So if you hear the word “computer” and there is a picture of a computer, touch the object gobbler as quickly as you can to give him the computer to put in his box.

But be careful! The mismatch gobbler likes to switch the pictures around. Anytime the word you hear does not match with the picture, touch the mismatch gobbler to give it back. So if you hear “red” and the picture is a BLACK motorcycle, touch the mismatch gobbler, because the color you heard did not match the color of the picture. Or, if you hear “computer” and the picture is a couch, touch the mismatch gobbler, because the word you heard was not the right name for the picture.

References

- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A.-M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, 87, 85–106. doi:10.1016/j.jecp.2003.10.002
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, 8, 71–82. doi:10.1076/chin.8.2.71.8724
- Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.
- Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, 49A, 5–28. doi:10.1080/027249896392784
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189–208. doi:10.1016/S0021-9924(03)00019-4

- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173. doi:10.1037/0033-295X.105.1.158
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition*. New York: Cambridge University Press.
- Blair, C. (2002). School readiness. Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*, 111–127.
- 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*, 647–663.
- Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology*, *28*, 561–571.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616. doi:10.1207/s15326942dn2802_3
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*, 282–298. doi:10.1111/j.1467-7687.2008.00675.x
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. doi:10.1111/1467-8624.00333
- Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*, 73–92. doi:10.1002/icd.298
- Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*, 672–691. doi:10.2307/1132197
- Denckla, M. B. (1996). A theory and model of executive function: A neuropsychological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 263–278). Baltimore, MD: Paul H. Brookes.
- Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, *95*, 465–481. doi:10.1037/0022-0663.95.3.465
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test: Manual* (4th ed.). Minneapolis, MN: Pearson Assessment.
- Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal working memory is involved in associative word learning unless visual codes are available. *Journal of Memory and Language*, *48*, 527–541. doi:10.1016/S0749-596X(02)00533-8
- Espy, K. A., Bull, R., Martin, J., & Stroup, W. (2006). Measuring the development of executive control with the Shape School. *Psychological Assessment*, *18*, 373–381. doi:10.1037/1040-3590.18.4.373
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Pathick, S., et al. (1993). *MacArthur–Bates Communicative Development Inventories: Gestures and words*. Baltimore: Paul H. Brookes.
- Frye, D., Zelazo, P. D., & Burack, J. A. (1998). Cognitive complexity and control: I. Theory of mind in typical and atypical development. *Current Directions in Psychological Science*, *7*, 116–121. doi:10.1111/1467-8721.ep10774754
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*, 1–16. doi:10.1002/acp.934
- Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language*, *59*, 267–333. doi:10.1006/brln.1997.1819
- Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology*, *79*, 192–227. doi:10.1006/jecp.2000.2586
- Hoyle, R. H., & Kenny, D. A. (1999). Sample size, reliability, and tests of statistical mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 195–222). Thousand Oaks, CA: Sage.
- Homer, B. D., & Nelson, K. (2009). Naming facilitates children's understanding of scale models: Language and the development of symbolic understanding. *Journal of Cognition and Development*, *10*, 115–134.
- Honomichl, R. D., & Chen, Z. (2011). Relations as rules: The role of attention in the Dimensional Change Card Sort task. *Developmental Psychology*, *47*, 50–60.
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*, 1447–1459.
- Jacques, S., & Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): A measure of executive function in preschoolers. *Developmental Neuropsychology*, *20*, 573–591. doi:10.1207/875656401753549807
- Jongejan, W., Verhoeven, L., & Siegel, L. S. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *Journal of Educational Psychology*, *99*, 835–851. doi:10.1037/0022-0663.99.4.835
- Jose, P. E. (2003). MedGraph-I: A programme to graphically depict mediation among three variables: The Internet version, version 2.0. Victoria University of Wellington, Wellington, New Zealand. Retrieved 13 June, 2011, from www.victoria.ac.nz/staff/paul-jose-files/medgraph/medgraph.php
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science*, *6*, 449–467. doi:10.1111/1467-7687.00300
- Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*, 677–692.
- Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*, 29–41. doi:10.1016/j.cognition.2009.08.014
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. doi:10.1146/annurev.psych.58.110405.085542
- Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental Science*, *12*, 1–18. doi:10.1111/j.1467-7687.2008.00754.x
- McBride-Chang, C., Zhou, Y., Cho, J.-R., Aram, D., Levin, I., & Tolchinsky, L. (2011). Visual spatial skill: A consequence of

- learning to read? *Journal of Experimental Child Psychology*, *109*, 256–262. doi:10.1016/j.jecp.2010.12.003
- McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007a). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*, 947–959.
- McClelland, M. M., Cameron, C. E., Wanless, S. B., & Murray, A. (2007b). Executive function, self-regulation, and social-emotional competence: Links to school readiness. In O. N. Saracho & B. Spodek (Eds.), *Contemporary perspectives on research in social learning in early childhood education* (pp. 83–107). Charlotte, NC: Information Age.
- Miller, S. E., & Marcovitch, S. (2011). Toddlers benefit from labeling on an executive function search task. *Journal of Experimental Child Psychology*, *108*, 580–592. doi:10.1016/j.jecp.2010.10.008
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167
- Müller, U., Zelazo, P. D., Lurye, L. E., & Liebermann, D. P. (2008). The effect of labeling on preschool children's performance in the Dimensional Change Card Sort Task. *Cognitive Development*, *23*, 395–408. doi:10.1016/j.cogdev.2008.06.001
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, *12*, 427–441. doi:10.1017/S0954579400003096
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. doi:10.3758/BRM.40.3.879
- Singer, B. D., & Bashir, A. S. (1999). What are executive functions and self-regulation and what do they have to do with language-learning disorders? *Language, Speech, and Hearing Services in Schools*, *30*, 265–273.
- Vygotsky, L. S. (1986). *Thought and language* (Rev. ed.; A. Kozulin, Ed.). Cambridge, MA: MIT Press. (Original work published 1934)
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing (CTOPP)*. San Antonio, TX: Pearson Education.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, *33*, 468–479. doi:10.1037/0012-1649.33.3.468
- Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, *22*, 306–317. doi:10.1037/a0018708
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*, 297–301. doi:10.1038/nprot.2006.46
- Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica*, *115*, 167–183.
- Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3, Serial No. 274), 1–119.