Journal of Experimental Child Psychology 192 (2020) 104779

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Brief Report

Between- and within-person contributions of simple reaction time to executive function skills in early childhood

Michael Willoughby [*], Yihua Hong, Kesha Hudson, Amanda Wylie

Education & Workforce Development, RTI International, Research Triangle Park, NC 27709, USA

ARTICLE INFO

Article history:
Received 13 November 2019
Revised 25 November 2019
26 November 2019
Available online 14 January 2020

Keywords:
Executive function
Simple reaction time
Processing speed
Early childhood
Within-person
Adaptive centering

ABSTRACT

This study tested whether the bivariate association between simple reaction time (SRT) and executive function (EF) performance that has been observed in early childhood represented a between- and/or within-person association. Up to three repeated assessments (i.e., fall, winter, and spring assessments from September to May) were available for 282 preschool-aged children (M_{age} = 4.2 years; 54% female) who participated in the Kids Activity and Learning Study. A series of three-level hierarchical linear models (repeated measures nested in child; child nested in classroom) was used to disaggregate the observed variation in EF and SRT into between-classroom, between-person, and within-person components. EF composite scores were regressed on two indicators of SRT, which reflected between- and within-child sources of variation, along with demographic covariates (child age, gender, and parental education). Both between-person (b = −21.2, p < 0.001) and within-person (b = −13.2, p < 0.001) sources of SRT variation were uniquely related to EF performance. These results are discussed with respect to interest in using SRT as a proxy for foundational cognitive processes that contribute to EF task performance in early childhood, including the appropriateness of using SRT to refine EF task scores.

---

[*] Corresponding author.
E-mail address: mwilloughby@rti.org (M. Willoughby).

https://doi.org/10.1016/j.jecp.2019.104779
0022-0965/

Introduction

Executive function (EF) skills are higher-order cognitive processes that regulate, bias, or otherwise influence lower-order foundational cognitive abilities (e.g., language, visual-spatial perception, joint attention, processing speed). These foundational cognitive abilities are rapidly developing in early childhood and may contribute to EF task performance (Espy, 2017). To the extent that children's performance on EF tasks reflects the contributions of EF skills *and* foundational cognitive abilities, EF task scores have ambiguous meaning.

Conventional wisdom holds that the measurement of EF can be improved by administering multiple tasks for each EF subdomain of interest and using factor analytic methods to define the subdomain of interest as the shared variation across tasks (e.g., Miyake et al., 2000). For example, inhibitory control could be defined by the variation that is shared across flanker, Stroop, and go/no-go tasks. An implicit assumption of this approach is that the idiosyncratic contributions that foundational cognitive abilities make to individual EF tasks are offset when multiple tasks are administered. However, as elaborated by Espy (2017), this assumption is untestable using standard factor analytic methods. To the extent that foundational cognitive abilities contribute to performance across a set of EF tasks, factor analytic approaches cannot differentiate EF from foundational cognitive processes. For example, if children's performances on flanker, Stroop, and go/no-go tasks all depend on inhibitory control, processing speed, and receptive language skills, a latent variable model that defines inhibitory control as the shared variance across tasks will not achieve the intended objective (i.e., the latent variable will conflate individual differences in processing speed and receptive language with inhibitory control).

One alternative to the conventional approach involves explicitly measuring foundational cognitive abilities that may contribute to EF task performance and using this information to refine the interpretation of EF task performance (Espy, 2017; van der Sluis, de Jong, & van der Leij, 2007). Willoughby, Blair, Kuhn, and Magnus (2018) recently described how a brief measure of simple reaction time (SRT), an index of processing speed, could be included in EF assessments and used to adjust EF task scores. In their study, the removal of SRT variation from individual EF task scores resulted in weaker associations between EF and chronological age as well as stronger associations between EF and poverty. In addition, EF task scores were less strongly correlated with each other after SRT-related variation was removed. The authors suggested that removing SRT-related variation from EF task scores may have practical and conceptual benefits. Practically, SRT measures can be obtained far more quickly than many other foundational cognitive processes that may contribute to EF task performance (e.g., crystallized knowledge related to task stimuli, receptive language skill that contributes to task comprehension). Conceptually, SRT may serve as a proxy for global aspects of cognitive development, including developmental variations in white matter tract development in the brain, that contribute to EF task performance (Chevalier et al., 2015; Scantlebury et al., 2014; Short et al., 2013, 2019; Treit, Chen, Rasmussen, & Beaulieu, 2014; Turken et al., 2008).

The speculation that SRT indexes individual white matter development implicitly assumed that there was a within-person association between SRT and EF (i.e., as an individual child's SRT improves, so does her or his EF task performance). However, this assumption was untested. The significant negative correlation that was observed between SRT and EF task performance may alternatively reflect a between-person association (i.e., individual differences in SRTs may simply be correlated with unmeasured factors or child attributes that are associated with EF task performance). Repeated-measures designs are required to differentiate between-person from within-person variation (Curran & Bauer, 2011). The primary objective of this study was to leverage repeated-measures data to test whether the negative association between SRT and EF performance (i.e., slower speed is associated with worse task performance) reflected between-person and/or within-person association. This distinction has implications for the appropriate interpretation of the association between SRT and EF as well as the utility of incorporating SRT tasks into EF assessments.

Method

Participants and procedures

A total of 282 children (54% female) were recruited from 71 classrooms in 13 preschools in the southeastern United States to participate in the Kids Activity and Learning Study. Participating children varied with respect to race (34% White, 31% Black or African American, 8% multiracial, 1% Asian, and 1% American Indian, with 17% of parents not reporting race) and ethnicity (8% Hispanic). Children ranged from 3 to 5 years of age ($M = 4.2$ years, $SD = 0.6$ at the fall assessment). Parental education was used to index family socioeconomic status, and 39% of parents reported attaining a 4-year college degree (or higher).

Children completed the same set of data collection activities in each of three 1-week data collection periods that occurred during the fall, winter, and spring seasons of an academic year (September–June). This study was focused exclusively on direct assessments of SRT and EF that occurred with a child assessor one morning per data collection period in a semiprivate space (e.g., quiet hallway, unused classroom). The RTI International institutional review board approved all study activities (Study No. 14239).

Measures

Executive function touch

EF Touch (Willoughby & Blair, 2016) is a computerized battery of six tasks that measure preschool-aged children's inhibitory control (three tasks), working memory (two tasks), and attention shifting (one task) skills as well as a brief measure of SRT (see https://www.youtube.com/channel/UC2g1F9-maborEHmZ-Frs-8rQ for an overview of EF Touch). Psychometric support for individual EF tasks and elaborated task descriptions have been provided elsewhere (Willoughby, Wirth, Blair, & Family Life Project Investigtors, 2012). Abbreviated task descriptions appear below. Note that each task in EF Touch was based on a cognitive paradigm or task that had previously been successfully used with young children.

Bubbles

This 30-item task measured SRT. A series of 30 bubbles of identical size, color, and shape appeared on the touch-screen monitor one at a time, and children were instructed to touch ("pop") each bubble as fast as they could. Items were presented for up to 5000 ms, and the time that transpired between stimuli onset and children's touch of the bubble was recorded. If an item was not touched, the item was considered inaccurate and the reaction time (RT) for that item was not recorded (>99% of items were accurately responded to at each assessment). Item responses that were faster than 400 ms were considered too fast to be plausible and were set to missing. The mean RT across all valid items was used as an index of SRT. To improve interpretability, SRT was scaled such that a 1-unit change referred to 1 s (not 1 ms).

Spatial conflict arrows

This 36-item spatial conflict task measured inhibitory control. Two "buttons" appeared on the left- and right-most sides of the touch-screen monitor. Arrow stimuli appeared sequentially either above the button to which they pointed (congruent items) or opposite of the button to which they pointed (incongruent items). Each arrow was presented for 3000 ms. The mean accuracy for incongruent items was used to index performance.

Silly sounds Stroop

This 17-item Stroop-like task measured inhibitory control. Each item displayed pictures of a dog and cat (the left–right placement on the screen varied across trials) and presented the sound of either a dog barking or a cat meowing. Children were instructed to touch the picture of the animal that did not make the sound (e.g., touch the cat when hearing a dog bark). Each item was presented for 3000 ms. The mean accuracy of all items was used to index performance.

Animal Go/No-Go

This 40-item go/no-go task measured inhibitory control. Individual pictures of animals were presented, and children were instructed to touch a centrally located "button" on their screen every time they saw an animal (the "go" response) except when that animal was a pig (the "no-go" response). Each item was presented for 3000 ms. The mean accuracy across all no-go responses was used to index task performance.

Working memory span

This 18-item span task measured working memory. Children were presented with arrays of one house, two houses, and three houses. Each house contained a picture of an outlined animal and a colored dot or a colored animal. Children verbally labeled the contents of each house. After a brief delay, the array of houses was displayed again without their contents. Children were asked to recall either the animal or color (of the animal) that was in each house (nonrecalled contents served as a distraction). The mean accuracy of responses was used to index task performance.

Pick the picture

This 32-item self-ordered pointing task measured working memory. Children were presented with arrays of two, three, four, and six pictures. For each picture set, children were initially instructed to touch any picture of their choice. On subsequent trials within that set, the pictures were presented in different locations, and children were instructed to pick a picture that had not yet been touched. The mean accuracy of responses was used to index task performance.

Something's the same

This 30-item task measured attention shifting. Children were presented with two pictures that were similar among the dimension(s) of color, shape, and/or size and were explicitly told how the two pictures are similar. Children were then presented with a third picture and asked to select one of the first two pictures that was the same as the new picture in some way. After 20 items, children were presented with all three pictures and were asked to select two pictures that were similar along one dimension and then two pictures that were similar along a different dimension. The mean accuracy of responses was used to index task performance.

EF touch scoring

Children completed an average of five or more of the six EF tasks that were administered at each assessment occasion (*M*s = 5.0, 5.3, and 5.5 tasks at the fall, winter, and spring assessments, respectively). Following earlier precedent and rationale (Willoughby, Blair, & Family Life Project Investigators, 2016; Willoughby, Kuhn, Blair, Samek, & List, 2017), children's performances across all completed EF tasks were combined to form an overall EF composite score. In addition, children's mean performance on the three inhibitory control tasks (Arrows, Silly Sounds Stroop, and Animal Go/No-Go) and two working memory tasks (Pick the Picture and Working Memory Span) were also combined to form inhibitory control and working memory composite scores.

Analytic approach

A series of three-level hierarchical linear models (where repeated measures were nested in children and children were nested in classroom) were used to investigate between- and within-person associations between children's SRT and EF task performance. We initially estimated an unconditional model for SRT and for each EF composite to characterize the proportion of variance that existed between classrooms, between persons, and within persons. Next, we regressed each of the EF composites on SRT and covariates, following the parameterization that was recommended by Wang and Maxwell (2015). The reduced form equation is presented below, with subscripts *i, j,* and *t* indexing person, classroom, and time (fall, winter, or spring), respectively.

$$EF_{ijt} = \beta_0 + \beta_1\left(SRT_{ijt} - \overline{SRT}_{ij}\right) + \beta_2\left(SRT_{ij} - \overline{SRT}_{j}\right) + \sum_{p=3}^{5}\beta_p\mathrm{Cov}_{ij} + r_j + u_{ij} + \varepsilon_{ijt}$$

Specifically, the contribution of SRT to EF was represented by both within- and between-person terms. The time-specific measure of SRT was centered at its person mean $\left(SRT_{ijt} - \overline{SRT}_{ij}\right)$, with β_1 representing a within-person effect; the person mean of SRT was centered at the grand mean $\left(SRT_{ij} - \overline{SRT}_{j}\right)$, with β_2 representing a between-person effect. We adjusted for child age, sex, and parental education level as covariates and included three variance components—r_j, u_{ij}, and ε_{ijt}—to estimate between-class, between-child, and within-child variation.

Results

Descriptive statistics

Descriptive statistics for the focal study variables are summarized in Table 1. Three points are noteworthy. First, children demonstrated decreases in SRT (they became faster) and increases in EF performance across the fall, winter, and spring assessments (~6 months). Second, consistent with our previous work, longer (slower) SRT was robustly negatively associated with EF task performance at each assessment occasion (*r*s = −0.28 to −0.51) as well as with children's average SRT across all measurement occasions (*r*s = −0.46 to −0.59). These associations indicated the potential existence of both within- and between-person effects. Third, although not tabled, child age was consistently correlated with both SRT and EF (|*r*|s = 0.45 to 0.60). Child age at the first assessment was included as a covariate to ensure that any associations between SRT and EF were not due to individual differences in age.

Table 1
Descriptive statistics.

Simple reaction time	*n*	*M* (*SD*)	
Fall	262	1.25 (0.32)	
Winter	264	1.09 (0.29)	
Spring	242	1.06 (0.27)	
Person average	282	1.14 (0.27)	
Executive function	***n***	***M* (*SD*)**	**Correlation with simple reaction time (*r*)**
Fall	262	59.09 (14.09)	−0.45
Winter	264	65.13 (14.31)	−0.51
Spring	242	68.57 (14.62)	−0.50
Person average	282	63.52 (13.14)	−0.59
Working memory			
Fall	252	56.31 (14.78)	−0.28
Winter	260	62.21 (12.94)	−0.38
Spring	234	64.37 (13.51)	−0.32
Person average	281	60.19 (11.46)	−0.46
Inhibitory control			
Fall	257	58.96 (21.11)	−0.41
Winter	261	65.50 (20.42)	−0.44
Spring	239	69.81 (20.29)	−0.46
Person average	281	63.57 (18.58)	−0.55

Note. Person average is the individual mean performance across available fall, winter, and spring assessments. Sample sizes vary due to differential requirements regarding which and how many individual tasks needed to be completed for a composite score to be created.

Hierarchical linear models

We estimated four unconditional models (one for each EF composite and one for SRT) to determine how much of the total observed variance for each variable was attributable to between-classroom, between-person, and within-person components. For the overall EF composite, 26%, 38%, and 36% of the total variation was attributable to between-classroom, between-person, and within-person differences, respectively. A similar pattern was evident for the measure of SRT (29%, 30%, and 40%, respectively) and the inhibitory control composite (19%, 39%, and 42%, respectively). The pattern for working memory differed in that a greater overall proportion of variance was attributable to within-person differences (12%, 25%, and 63% attributable to between-classroom, between-person, and within-person differences, respectively). These descriptive results indicated that appreciable variation existed across and within children for both the focal predictor and the EF outcomes, supporting our ability to make meaningful tests of study aims.

Next, we estimated three conditional models. For each outcome, both between- and within-person effects were statistically significant (see Table 2). Specifically, the within-person effects indicated that every 1-s improvement (reduction) in SRT across measurement occasions would be associated with an approximately 13%, 11%, and 15% improvement in overall EF, working memory, and inhibitory control composite scores, respectively (all *p*s < 0.0001). The between-person effects indicated that for every 1 s faster that children's average SRT was relative to the sample mean, children demonstrated approximately 21%, 13%, and 28% improvements in overall EF, working memory, and inhibitory control composites, respectively (all *p*s < 0.0001).

Discussion

This study replicated and extended previous work that related individual differences in SRT to EF task performance (i.e., slower performance on the SRT task was associated with poorer performance on EF tasks). Across measurement occasions, instances of improved (faster) SRT were associated with corresponding improvements in EF task performance. These within-person associations were evident even after controlling for individual differences in SRT (person mean averages), as well as demographic characteristics, including chronological age, which is correlated with both SRT and EF. These results are consistent with the idea that children's processing speed is a foundational cognitive process that contributes to EF task performance and that may index white matter tract development.

Significant between-person effects were also evident. Children who on average demonstrated overall faster performance on the SRT task also performed better on EF tasks. This association is consistent with the idea of person-level confounding. Stable individual differences in SRT may correlate with

Table 2
Estimated within- and between-person effects of SRT.

	Executive function		
	Executive function composite	Working memory	Inhibitory control
Fixed effect	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Intercept	63.88 (0.66)***	60.54 (0.59)***	64.13 (0.91)***
Repeated SRT	−13.22 (1.95)***	−11.26 (2.57)***	−15.26 (3.14)***
Person mean SRT	−21.17 (2.72)***	−13.46 (2.73)***	−27.76 (4.04)***
Random effect			
Level 1 (repeated level)	73.42***	122.45***	181.31***
Level 2 (person level)	58.58***	43.56***	124.92***
Level 3 (class level)	6.86	0.93	6.77
−2LL	5804.3	5875.8	6365.2
Akaike information criterion	5810.3	5881.8	6371.2
Bayesian information criterion	5817.2	5888.7	6378.2

Note. SRT, simple reaction time.
*** $p < 0.001$.

(and, hence, be a proxy for) a multitude of genetic and/or experiential factors that are related to EF task performance. For example, individual differences in children's SRT may reflect prenatal differences in nutrition (Brouwer-Brolsma, Vrijkotte, & Feskens, 2018), toxicant exposure (Derauf et al., 2012), or preterm birth (Rose, Feldman, & Jankowski, 2009), all of which have been related to the development of processing speed.

Taken together, these results provide mixed support for adjusting young children's EF task scores based on their SRT performance. Although SRT contributed to individual-level performance on EF tasks, it also correlated with unmeasured factors that are associated with EF task performance. In studies that involve repeated assessments of EF, adjusting EF task scores for within-person contributions of SRT can be accomplished using the approach that we demonstrated here. However, in studies that involve a single assessment, EF task scores that have been adjusted for SRT have uncertain interpretation. Dennis et al. (2009) elaborated the problems associated with using IQ as a covariate in studies involving cognitive outcomes. We believe that many of these same problems apply to using SRT as a covariate for EF tasks in circumstances when the between- and within-person components of SRT cannot be distinguished.

In studies that involve a single assessment of EF, it may be preferable to use tasks that permit the simultaneous measurement of SRT and EF ability (rather than a separate assessment of SRT). In early childhood, EF task performance is virtually always determined based on the accuracy of children's performance. However, EF tasks that are used with older children and adults routinely include baseline or reference items that index SRT as well as test items that putatively measure EF processes (e.g., go vs. no-go trials on a go/no-go task, congruent vs. incongruent items on a flanker task). Taking the difference in RT between reference and test items is a common strategy for inferring EF performance in a way that adjusts for the contributions of SRT (it also avoids ceiling effects that would be evident if only the accuracy of responding was considered). It is an open question how best to leverage accuracy and RT data to infer EF task performance and SRT among preschool-aged children (see Magnus, Willoughby, Blair, & Kuhn, 2019).

A few limitations and caveats are noteworthy. First, we scaled between- and within-person predictors in whole seconds to facilitate ease of interpretation. A 1-s interval is arguably too large to represent the expected magnitude of within- or between-person changes in SRT. Although the scaling time to tenths of seconds would reduce the magnitude of estimated effects, it would not influence our conclusions. Second, although it is tempting to compare the magnitudes of between- and within-person effects, we would advise against doing so. It is not clear that the expected variation in changes in SRT within a child is comparable to the variation in SRT that is observed across children. Third, the degree of within-person variation that is evident in SRT and EF is contingent on the age of children being studied as well as the span of time during which data are collected. It is not clear whether these results would generalize to older children or to studies that measured SRT and EF over appreciably different periods of time. Fourth, although we motivated this study by an interest in foundational cognitive processes that contribute to EF task performance, we have limited our attention to SRT. A broader consideration of other foundational cognitive processes that contribute to EF task performance is clearly merited.

Given widespread interest in the measurement of EF skills in early childhood, it is imperative that the field develops assessments that both are easy to administer and have unambiguous interpretation. Although ease of use is no longer an issue, ambiguities in the interpretation of EF scores remain. The ability to measure dynamic EF processes in early childhood is complicated by the fact that many of the foundational cognitive abilities that EFs act on are themselves rapidly developing and may contribute to EF task performance. Despite the conceptual appeal of using brief measures of SRT as a proxy for foundational cognitive processes that contribute to EF performance, the results of this study underscore that this approach is most appropriate for studies involving repeated measures (where between- and within-person sources of variation can be distinguished). More creative solutions are required for distinguishing EF from SRT-related variation when assessments are limited to a single measurement occasion.

Acknowledgment

This study was supported by an Institute of Educational Sciences grant (R305A160035). The views expressed in this manuscript are those of the authors and do not necessarily represent the opinions and positions of the funders.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2019.104779.

References

Brouwer-Brolsma, E. M., Vrijkotte, T. G. M., & Feskens, E. J. M. (2018). Maternal vitamin D concentrations are associated with faster childhood reaction time and response speed, but not with motor fluency and flexibility, at the age of 5–6 years: The Amsterdam born children and their development (ABCD) study. *British Journal of Nutrition, 120*, 345–352.

Chevalier, N., Kurth, S., Doucette, M. R., Wiseheart, M., Deoni, S. C., Dean, D. C., 3rd, ... LeBourgeois, M. K. (2015). Myelination is associated with processing speed in early childhood: Preliminary insights. *PLoS ONE, 10*(10). e139897.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*, 583–619.

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society, 15*, 331–343.

Derauf, C., Lester, B. M., Neyzi, N., Kekatpure, M., Gracia, L., Davis, J., ... Kosofsky, B. (2012). Subcortical and cortical structural central nervous system changes and attention processing deficits in preschool-aged children with prenatal methamphetamine and tobacco exposure. *Developmental Neuroscience, 34*, 327–341.

Espy, K. (2017). The changing nature of executive control in preschool. *Monographs of the Society for Research in Child Development, 81*(4), 1–179.

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*, 1296–1306.

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

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2009). Information processing in toddlers: Continuity from infancy and persistence of preterm deficits. *Intelligence, 37*, 311–320.

Scantlebury, N., Cunningham, T., Dockstader, C., Laughlin, S., Gaetz, W., Rockel, C., ... Mabbott, D. (2014). Relations between white matter maturation and reaction time in childhood. *Journal of the International Neuropsychological Society, 20*, 99–112.

Short, S. J., Elison, J. T., Goldman, B. D., Styner, M., Gu, H. B., Connelly, M., ... Gilmore, J. H. (2013). Associations between white matter microstructure and infants' working memory. *NeuroImage, 64*, 156–166.

Short, S. J., Willoughby, M. T., Camerota, M., Stephens, R. L., Steiner, R. J., Styner, M., & Gilmore, J. H. (2019). Individual differences in neonatal white matter are associated with executive function at 3 years of age. *Brain Structure and Function, 224*, 3159–3169.

Treit, S., Chen, Z., Rasmussen, C., & Beaulieu, C. (2014). White matter correlates of cognitive inhibition during development: A diffusion tensor imaging study. *Neuroscience, 276*, 87–97.

Turken, A., Whitfield-Gabrieli, S., Bammer, R., Baldo, J. V., Dronkers, N. F., & Gabrieli, J. D. (2008). Cognitive processing speed and the structure of white matter pathways: Convergent evidence from normal variation and lesion studies. *NeuroImage, 42*, 1032–1044.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35*, 427–449.

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*, 63–83.

Willoughby, M. T., & Blair, C. B. (2016). Longitudinal measurement of executive function in preschoolers. In J. Griffin, L. Freund, & P. McCardle (Eds.), *Executive function in preschool age children: Integrating measurement, neurodevelopment and translational research* (pp. 91–113). Washington, DC: American Psychological Association Press.

Willoughby, M. T., Blair, C. B., & Family Life Project Investigators (2016). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment, 28*, 319–330.

Willoughby, M. T., Blair, C. B., Kuhn, L. J., & Magnus, B. E. (2018). The benefits of adding a brief measure of simple reaction time to the assessment of executive function skills in early childhood. *Journal of Experimental Child Psychology, 170*, 30–44.

Willoughby, M. T., Kuhn, L. J., Blair, C. B., Samek, A., & List, J. A. (2017). The test–retest reliability of the latent construct of executive function depends on whether tasks are represented as formative or reflective indicators. *Child Neuropsychology, 23*, 822–837.

Willoughby, M. T., Wirth, R. J., Blair, C. B., & Family Life Project Investigators (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment, 24*, 418–431.