

Peers Matter: Links Between Classmates' and Individual Students' Executive Functions in Elementary School

Jenna E. Finch 

University of Nebraska–Lincoln

Elisa B. Garcia

SRI International

Michael J. Sulik

Jelena Obradović

Stanford University

Students' executive functions (EFs) are linked to school success. Although school-age children spend much of their time interacting with peers, few studies have explored how children's classmates may promote EF development in elementary school. In this study, we test whether mean levels and variability in classmates' EF skills are associated with growth in individual students' accuracy and speed on EF tasks among third, fourth, and fifth graders (N = 806). We find that classmates' speed, but not accuracy, on EF tasks is linked to significant improvements in individual students' EFs over the school year. Classmates' average EFs, as indexed by faster accurate responses on EF tasks, are associated with improvements in individual students' speed on EF tasks. These results were robust to the inclusion of individual students' general processing speed. In contrast, variability in classmates' accuracy and speed on EF tasks was not associated with individual students' EF growth. Our results highlight the role of peers and the school context for EF development in middle childhood.

Keywords: executive functions, peer effects, middle childhood, elementary school, self-regulation

EXECUTIVE functions (EFs) are a set of higher-order cognitive skills that help students regulate their behaviors and attention (Diamond, 2013). Children's EFs are important for the acquisition of academic skills (Cartwright, 2012; Kolkman, Hoijsink, Kroesbergen, & Leseman, 2013) and are predictive of academic achievement over and above intelligence and socioeconomic status (Blair & Razza, 2007; Bull, Espy, Wiebe, Sheffield, & Nelson, 2011). Students engage their EFs to behave appropriately in the classroom context, to learn new academic content, and to play and collaborate with peers (Cairano, Visu-Petra, & Settanni, 2007; Finch & Obradović, 2017; Sasser, Bierman, & Heinrichs, 2015). During elementary school, interactions with peers become an increasingly important way for children to practice regulating their attention and behaviors (Gifford-Smith & Brownell, 2003). A few studies explored the role of classmates' EF skills for EF development in early childhood (Montroy, Bowles, & Skibbe, 2016; Skibbe, Phillips, Day, Brophy-Herb, & Connor, 2012; Weiland & Yoshikawa, 2014). The current study extends previous literature to examine whether classmates' EFs are associated with the development of individual students' EF growth in

middle childhood. We explore how average levels and variability in classmates' performance on direct tests of EFs relate to changes in individual students' EF skills during the school year.

EFs in the Elementary School Context

Although school influences on the development of EF skills have been predominantly studied in the early childhood period (Montroy et al., 2016; Weiland, Ulvestad, Sachs, & Yoshikawa, 2013), EFs continue to develop through the elementary school years and beyond (Best, Miller, & Jones, 2009; Lee, Bull, & Ho, 2013). In middle childhood, students face increased attentional and cognitive demands, as they are asked to manage their behaviors with less direct adult scaffolding. During this developmental period, students are expected to independently listen and keep track of directions, collaborate on group activities, complete work, and play appropriately with peers—all of which require children to regulate their attention and behaviors (Finch & Obradović, 2017; Sasser et al., 2015).



In preschool, EFs are typically measured by children's accuracy (i.e., percentage correct) on tasks. As children get older, they demonstrate ceiling effects on accuracy measures; as such, researchers have also included children's reaction times (RTs) on EF tasks such that faster RTs on EF-demanding task trials indicate better EFs (Zelazo, Blair, & Willoughby, 2016). In contrast, children's RTs on task trials where EF demands are absent represent children's general processing speed. During middle childhood, children's abilities to monitor their performance and adjust their behavior (i.e., metacognition) increase substantially, supporting improvements in EFs (Best & Miller, 2010). For example, children begin to show a "speed-accuracy tradeoff" by slowing down to increase their accuracy during this developmental period (Davidson, Amso, Anderson, & Diamond, 2006). Thus, it is important to examine how children's accuracy and RTs on EF tasks are each related to student outcomes during middle childhood.

In middle childhood, children engage in more frequent peer-to-peer interactions, which are less likely to be mediated by adults (Gifford-Smith & Brownell, 2003). The quality of children's social interactions with peers plays a key role in shaping their social-emotional and academic skills in middle childhood (Pedersen, Vitaro, Barker, & Borge, 2007; Véronneau, Vitaro, Brendgen, Dishion, & Tremblay, 2010). Peer interactions become increasingly complex during middle childhood, requiring students to regulate their behaviors to engage in organized play and discussions. Indeed, there are reciprocal associations between developing social skills and children's EFs across elementary and middle school (Holmes, Kim-Spoon, & Deater-Deckard, 2016). Therefore, it is possible that interacting with peers with high EFs may provide children with examples of well-regulated behavior, more opportunities to practice self-regulation skills, and the motivation to regulate themselves.

Furthermore, well-regulated peers might create an overall classroom environment conducive to practicing self-regulation skills. One study, for example, found that children who were better regulated spent less class time on unproductive activities, such as waiting, off-task behaviors, or disruptions (Day, Connor, & McClelland, 2015); others studies showed that children who were engaged in classroom tasks gained more self-regulation skills during the school year (Bierman, Torres, Domitrovich, Welsh, & Gest, 2009; Williford, Vick Whittaker, Vitiello, & Downer, 2013). The classroom composition of peers' EFs likely shapes the culture of the classroom environment and affects classroom norms and expectations for individual students' self-regulation.

Average Peer Effects and Children's Development

Most studies examining the role of classmates' skills for children's development focused on links between classroom levels of children's academic achievement or behavioral

problems and individual students' development in those domains (Gottfried, 2014; Hanushek, Kain, Markman, & Rivkin, 2003; Thomas, Bierman, & Powers, 2011). Typically, researchers estimate "peer effects" by analyzing how classmates' average levels of a skill in the fall are linked to growth in individual students' skills from fall to spring. Having classmates with higher average levels of reading and math scores has been associated with growth in individual students' test scores during elementary school, even after controlling for a robust set of child-, classroom-, and teacher-level covariates (Boucher, Bramoullé, Djebbari, & Fortin, 2014; Gottfried, 2014; Hanushek et al., 2003). The literature also supports similar peer effects with regard to students' behavior problems such that being in classrooms where peers have high average levels of aggression and delinquency is linked to increases in individual students' behavior problems and reductions in individual students' test scores in elementary and middle school (Figlio, 2005; Mercer, McMillen, & DeRosier, 2009; Muller, Hofmann, Fleischli, & Studer, 2016; Thomas et al., 2011; Yudron, Jones, & Raver, 2014). Given that EFs are related to both academic achievement and behavioral problems (Bull et al., 2011; Ciairano et al., 2007; Jacob & Parkinson, 2015; Kolkman et al., 2013; Schoemaker, Mulder, Deković, & Matthys, 2013), classmates' EFs might be similarly important for children's development during middle childhood.

A few studies explored associations between classmates' average levels of EFs and children's cognitive skills in the early childhood period. There is some evidence that classmates' EFs are linked to preschoolers' growth in early language and literacy skills (Montroy et al., 2016; Skibbe et al., 2012), although one study did not find an association between classmates' EFs and preschoolers' vocabulary achievement (Weiland & Yoshikawa, 2014). To our knowledge, only two studies examined whether classmates' EFs are linked to individual children's growth in EFs across a single academic year. Montroy and colleagues (2016) found that classmates' average levels of EFs were positively linked to growth in individual preschoolers' EFs in a relatively advantaged sample of children. In contrast, Weiland and Yoshikawa (2014) did not find evidence that classmates' average levels of EFs were linked to individual children's growth in EFs in a diverse sample of predominantly low-income preschoolers.

These divergent findings may be due to many factors, including sample characteristics and study design. Montroy et al. (2016) drew participants from private and public community-based preschool settings that did not have specific curricula. Children in the sample were predominantly Caucasian and had relatively well-educated mothers. In contrast, Weiland and Yoshikawa (2014) drew participants from urban public prekindergarten programs that were part of a large-scale intervention study to improve preschool quality. The intervention program, which increased students' EFs

(Weiland & Yoshikawa, 2013), included empirically validated curricula with personalized coaching for teachers. The sample was racially and ethnically diverse, and most children were eligible for free or reduced-price lunch. It is possible that differences between community-based child care centers and more public school-based prekindergarten programs influence peer effects, but more information is needed about why classmates' EFs may be more salient in certain school contexts.

In addition, coverage (i.e., the number of children used to estimate the classroom-level average EFs) differed between these studies. Similar to authors of other peer effect studies (Justice, Petscher, Schatschneider, & Mashburn, 2011; Mashburn, Justice, Downer, & Pianta, 2009), Weiland and Yoshikawa (2014) randomly tested five students per classroom on average. In contrast, Montroy et al. (2016) tested 12 students per classroom on average, with over half of the students recruited in most (65%) classrooms. Higher classroom-level coverage supports more precise estimate of classmates' skills and may have enabled Montroy and colleagues to better detect peer effects.

Variability of Classmates' Skills

To more fully characterize the distribution of classmates' skills, some researchers measured the variability of classmates' skills, in addition to average levels of classmates' skills. The standard deviation of classmates' skills represents how closely individual students' skills are clustered around the classroom mean. Theoretically, classrooms that are more homogeneous in students' skills might run more smoothly. Classroom activities would likely be at the appropriate level to support most students' academic and behavioral development when students have relatively similar skill levels (Tomlinson et al., 2003). Two studies that explored variability in students' academic skills found null effects on elementary school students' academic growth, based on state-level administrative data with high coverage (Burke & Sass, 2013; Hanushek et al., 2003). This suggests that variability in classmates' achievement may not be strongly linked to students' academic learning in the early grades. However, variability in classmates' social-emotional skills might operate differently in the classroom context.

Greater levels of variability in peers' skills may not necessarily have a negative effect on individual students' socio-emotional development. For example, research on parent-child interactions highlights the benefits of variability in emotional states and switching more frequently among different emotional states during an interaction, including negative emotions (Granic & Patterson, 2006; Hollenstein, 2007). These studies suggest that more variability in parents' and children's emotional states may help children practice regulating positive and negative emotions by providing them opportunities to deal with a variety of emotional states

(Granic & Patterson, 2006; Hollenstein, Granic, Stoolmiller, & Snyder, 2004). Similar mechanisms might extrapolate to peer groups such that more variability in peers' self-regulation skills would provide individual children with more practice regulating their emotions and behaviors in response to their peers' actions.

Empirical studies examining the effects of variability in classmates' nonacademic skills on individual students' development are rare. The main obstacle is that there are few direct measures of children's nonacademic skills (e.g., EFs) that can be easily used in a classroom context at scale (McKown, 2017). Using teacher reports of children's mental health symptoms, Yudron and colleagues (2014) demonstrated that more variability in classmates' externalizing behavior problems was linked to increased internalizing behavior problems and lower social competence for preschoolers. No studies, to our knowledge, examined whether variability in classmates' EFs are associated with individual students' EF growth.

Current Study

This study is the first to examine associations between classmates' EFs and individual children's EF development in middle childhood. We leveraged a unique data set that includes repeated direct assessments of third-, fourth-, and fifth-grade students' EFs for nearly all children across 33 classrooms. This study design enabled us to have high coverage of students within each classroom ($M = 90\%$). We used accuracy and RTs across four tasks to create two measures of classmates' EFs: average levels and variability. First, we investigated how average levels of classmates' EFs are linked to growth in students' EF skills from fall to spring. We hypothesized that higher average levels of classmates' EFs would be linked to more growth in individual students' EFs between the fall and spring of the school year. Second, we conducted exploratory analyses to examine whether variability in classmates' EFs are linked to growth in individual students' EF skills, after controlling for average levels of classmates' EFs.

Method

Sample and Setting

Third-, fourth-, and fifth-grade teachers were recruited from two public school districts in the San Francisco Bay Area in the academic year before data collection took place. All students in target classrooms were invited to participate through home mailings and Back to School Night presentations. For children's participation in classroom assessments of EF skills, parents provided passive consent (i.e., parents were notified of the study procedures and were offered an opportunity to withdraw from the study; only one family withdrew its child from participation in this part of the

study), and each student provided an active verbal assent ($n = 813$). Furthermore, we obtained an active written consent from a subset of 569 parents to access students' school administrative data, including state-administered standardized test scores and demographic data about parents and children. All 33 teachers gave informed consent.

The analytic sample for this study included 806 children (275 third graders, 354 fourth graders, and 177 fifth graders; 48% female) with valid data for the fall or spring EF assessment. A quarter of students in our sample (24.69%) were in multigrade classrooms, where students from different grade levels were combined into a single classroom with one teacher (e.g., fourth- and fifth-grade class). Some third-grade students were in combined second- and third-grade classrooms. Second-grade students in these classrooms were not included in the study. Our sample was socioeconomically and ethnically diverse: Among the 70% of parents who reported their children's ethnicity, 6% were identified as African American, 23% as Caucasian or White, 34% as Asian or Pacific Islander, 32% as Hispanic/Latino, and 6% as multiracial or other. Among the 58% of parents who reported their educational attainment, 37% had a high school degree or less, and 21% had a graduate degree.

Procedures

Data for this study were collected from two assessments: in the fall and spring of a single academic year. The fall and spring assessments were collected within 3 weeks of the beginning and end of the school year, respectively. All students completed EF tasks on tablet computers at each time point. Developmentally appropriate and widely used EF tasks were adapted to Android tablet computers. During the assessment, each child was given a tablet computer, and all students in the classroom completed the four EF tasks as a group (Obradović, Sulik, Finch, & Tirado-Strayer, 2018). In each session, three research assistants were present to administer the EF tasks to the entire class at one time. One assessor stood at the front of the class, using large posters to explain the task and guide children through each step. The two remaining assessors assisted children with their tablets and helped address any technical problems as needed.

The tasks and instruction screens were illustrated with cartoon pictures, designed to be fun and appealing to children. Tasks were presented with simplified rules that children could read on their own. Each task included a set number of practice items and test trials. Password-locked screens were used to ensure that all students were on the correct task when the research assistant explained the instructions and that they began each task at the same time. Except for the Digit Span Backward task, each set of EF trials was timed so that children finished at approximately the same time. Findings from a prior study demonstrated that this group EF assessment administered with tablets was a valid

approach to directly measuring students' EFs (Obradović et al., 2018).

Measures

Executive functions. The Digit Span Backward task, drawn from the Wechsler Intelligence Scale for Children–IV (Flanagan & Kaufman, 2009), is a measure of children's working memory skills. A series of digits were presented sequentially on the tablet screen. The student was instructed to enter those numbers backward onto a numeric keypad after the last digit was presented. There were four practice trials, each with strings that were two digits long. These practice trials were followed by eight test trials of increasing difficulty (two trials each of two, three, four, and five digits). Accuracy scores were computed as the proportion of correct test trials (fall, $\alpha = .80$; spring, $\alpha = .77$).

The Hearts and Flowers (HF) task is a measure of inhibitory control and cognitive flexibility skills (Davidson et al., 2006). There were three blocks: 12 congruent "heart" trials, 12 incongruent "flower" trials, and 33 mixed "heart and flower" trials. On each trial, students were presented with an image of a heart or a flower, which appeared on the right or left side of the screen. For congruent heart trials, students were instructed to press the button on the same side as the heart. For incongruent flower trials, students were instructed to press the button on the opposite side of the flower. Accuracy scores were drawn from the incongruent block (fall, $\alpha = .89$; spring, $\alpha = .83$) and the mixed block (fall, $\alpha = .90$; spring, $\alpha = .90$). Although the window of time in which children could respond to the HF task (750 ms) was based on previous research (Davidson et al., 2006), the pacing for the mixed block was too rapid for children in this study, resulting in many missing RT scores during the mixed block. Consequently, RT scores were drawn only from the incongruent block (fall, $\alpha = .87$; spring, $\alpha = .94$), and we increased the maximum response time to 1,250 ms for the spring assessment.

The Multi-Source Interference Test (MSIT) is a measure of inhibitory control skills (Bush & Shin, 2006). There were two blocks: 24 congruent trials and 24 incongruent trials. In both blocks, students were presented with a sequence of three digits. For each trial, two of these digits (the distractors) were the same, and one differed (the target; e.g., "2 2 1"). Students were instructed to press a button whose numeric value corresponded to that of the target. For example, the correct response to the sequence "2 2 1" would be "1." For the congruent trials, the distractors were always zeroes, and the position of the target always corresponded to the numeric value of the correct button press (i.e., "1 0 0," "0 2 0," "0 0 3"). For the incongruent trials, the distractors were nonzero, and the numeric value of the correct button press was always different from the position of the correct response (e.g., "2 3 3," "2 2 1," "1 3 1"). Accuracy and RT scores from the incongruent

block of trials were used (fall accuracy, $\alpha = .94$; spring accuracy, $\alpha = .92$; fall RT, $\alpha = .89$; spring RT, $\alpha = .91$).

The Flanker task is a measure of students' inhibitory control and cognitive flexibility skills (Diamond, Barnett, Thomas, & Munro, 2007). There were three blocks: 17 blue fish trials, 17 pink fish trials, and 45 mixed trials (blue and pink fish). Students were asked to focus on a given stimulus while inhibiting attention to stimuli flanking it. They were shown a row of fish on a screen and told to press the right or left arrow, depending on the direction that the target fish was facing. In the first block (blue fish), the target was the middle fish, whereas in the second block (pink fish), the target fish were the flanking outside fish. In the third block, pink and blue fish were mixed. During congruent trials, all fish faced the same direction, whereas during incongruent trials, the middle and outside fish did not face the same direction. Accuracy scores were drawn from the 12 incongruent blue and pink trials (fall accuracy, $\alpha = .96$; spring accuracy, $\alpha = .97$) and 45 mixed trials (fall accuracy, $\alpha = .89$; spring accuracy, $\alpha = .91$). We found that children who were slower on the Flanker mixed block demonstrated higher accuracy on it and better scores on the other task blocks. This suggests that quicker RTs during the Flanker mixed block are not a valid measure of EFs; therefore, RT scores were drawn from the 12 incongruent blue and pink trials only (fall RT, $\alpha = .94$; spring RT, $\alpha = .95$).

Scoring of EF tasks. Anticipatory responses—defined as a response <200 ms after stimulus presentation—were recoded as missing for the accuracy and RT scores. Furthermore, the HF, MSIT, and Flanker tasks were timed such that students were unable to respond after 750 ms (fall) / 1,250 ms (spring), 2,500 ms, and 1,500 ms, respectively. If the student failed to respond during this window, the trial was counted as incorrect for the accuracy score and as missing for the RT score. Finally, following standard practice, RT scores were calculated only for the accurate trials and not for the first trial in each block. Outliers, defined as accuracy or RT scores that were >4 *SD* above or below the sample mean, were winsorized to the highest nonoutlier value observed for that task. Across all task blocks, there was only one outlier value in the fall and four in the spring.

We used confirmatory factor analysis to inform our data reduction approach for the accuracy and RT scores. Mplus 7.3 (Muthén & Muthén, 2014) was used to estimate these models. For the accuracy composite, we started with six indicator variables: Digit Span Backward (working memory), MSIT incongruent block (inhibitory control), HF incongruent block, HF mixed block (cognitive flexibility), Flanker blue and pink incongruent blocks (inhibitory control), and Flanker mixed block (cognitive flexibility). To account for similarity between different blocks for each task, we estimated a one-factor model in which we correlated the residual variances for the two HF blocks and for the two

Flanker blocks. Model fit was acceptable in both fall and spring—fall: $\chi^2(df=7) = 26.529, p < .001$, root mean square error of approximation = .062, comparative fit index = .964, standardized root mean square residual = .033; spring: $\chi^2(df=7) = 20.802, p = .004$, root mean square error of approximation = .052, comparative fit index = .980, standardized root mean square residual = .026. Given these confirmatory factor analysis results, we standardized and averaged the HF incongruent and mixed blocks and the Flanker incongruent and mixed blocks to create a single score for each task. We then averaged and standardized the scores from each of the four tasks to create a single accuracy composite score for the fall and spring (fall, $\alpha = .62$; spring, $\alpha = .64$). The spring task scores were standardized per the fall mean and standard deviation to measure growth in EF scores from fall to spring. As shown here, fall-standardized spring scores (z_{is}) were calculated as individual student i 's raw scores on each task (x_{is}) in the spring minus the fall mean score on the task (x_f) and divided by the sample fall standard deviation on the task (s_f). Standardizing spring task scores based on the fall mean and standard deviation allowed us to measure growth in EF scores from fall to spring.

$$z_{is} = \frac{x_{is} - x'_f}{s_f}$$

For the RT composite, we started with three indicator variables: MSIT incongruent block (inhibitory control), HF incongruent block (inhibitory control), and Flanker blue and pink incongruent blocks (inhibitory control). We examined the normality of the RT variables using the skewness and excess kurtosis indices (see Appendix Table A1). All variables were relatively normally distributed; thus, we did not need to transform the variables. We standardized and averaged the scores from each of the three tasks to create a single RT composite score for the fall and spring (fall, $\alpha = .57$; spring, $\alpha = .72$). As with the accuracy composite score, the spring RT task scores were standardized with the fall mean and standard deviation. For ease of interpretation, RT composite scores were reversed so that higher scores can be interpreted as children demonstrating better EF skills.

We calculated difference scores by subtracting the average RT on congruent trials (which have no EF demands) from the average RT on incongruent or mixed trials (which do have EF demands). Across all tasks, the difference scores were not significantly correlated with children's accuracy composites on the same blocks or with measures of their academic achievement. Therefore, we did not use the difference scores, as they did not seem to capture EFs in our sample.

A measure of processing speed was created from three indicator variables: MSIT congruent block, HF congruent block, and Flanker blue congruent trials blocks. We standardized and averaged the scores from each of the three

tasks to create a single processing speed composite score for the fall ($\alpha = .73$). For ease of interpretation, the processing speed composite scores were reversed so that higher scores can be interpreted as children demonstrating quicker processing speed.

Classmates' EFs. We created two measures of classmates' EFs from students' individual EF composites. We used leave-out measures of classmates' EFs so that each student's own score was not included in the calculation of the composition of classmates' EFs. First, the leave-out ($N - 1$) mean represented the average levels of EFs for a student's classmates. Second, the leave-out ($N - 1$) *SD* around each classroom's mean represented the amount of heterogeneity of peers' EFs at the classroom level.

The average number of children whose EFs were assessed in the fall was approximately 23 per classroom ($SD = 5.32$, range: 12–32). This represented, on average, 90% ($SD = 7\%$, range: 76%–100%) of students in the class, which far exceeds previous peer effect studies that measured classmates' skills with a quarter to half of the children in classrooms (Justice et al., 2011; Mashburn et al., 2009; Montroy et al., 2016; Weiland & Yoshikawa, 2014).

Covariates. Child gender (0 = male, 1 = female) was included as a covariate because of prior research demonstrating gender differences in students' EF scores (Matthews, Ponitz, & Morrison, 2009). Student age was included because children show significant age-related gains in EF skills during elementary school (Lee et al., 2013). Class average grade was included to account for the mixed-grade classrooms in the study. Inherently, mixed-grade classrooms will have more variability in EFs, as students have more variability in age. Finally, a school-level measure of the proportion of free and reduced-priced lunch was included to control for school-level differences in student disadvantage. No other covariates were tested and removed.

Analytic Plan

To explore how classmates' EFs were linked to individual students' EF growth, we estimated regression models in a hierarchical linear modeling framework (also called *multi-level modeling*) such that students were at Level 1 and classrooms were at Level 2. To adjust parameter estimates and standard errors for the clustering of students within classrooms, all models included random intercepts for classrooms. In Models 1a and 2a, we predicted students' spring accuracy and RT scores, respectively, with the leave-out mean, controlling for fall accuracy and RT scores (respectively), student gender, classroom average age, and the proportion of students receiving free or reduced-price lunch in the school. Model 1b adds a measure of fall RT when predicting spring accuracy, and Model 2b adds a measure of fall

accuracy when predicting spring RT. Models 1c and 2c add a measure of fall processing speed. To each of these models, we added a measure of the standard deviation of EFs in the classroom, to explore whether variability in EFs explained growth in individual students' EFs.

Missing data were addressed with multiple imputation with 20 complete data sets, which is methodically superior than other methods of addressing missing data in complex data sets (Enders, 2017). The percentage of missing data for all study variables ranged from 0.5% for student gender to 16% for the accuracy score on the fall Digit Span Backward task. The accuracy and RT scores for EF tasks were first imputed and then standardized and averaged to create the accuracy and RT composite scores, as described earlier. There were no differences in rates of missing data for fall EF tasks by student age, grade, gender, or school-level proportion of students receiving free or reduced-price lunch. In the spring, students from schools with higher proportions of students receiving free or reduced-price lunch were more likely to have missing data across all four tasks. Furthermore, younger students were more likely to have missing data in the spring on the HF and Flanker tasks.

Results

Descriptive Statistics and Bivariate Correlations

Table 1 shows that students' individual EF scores, on both RT and accuracy measures, increased between the fall and spring (fall accuracy = 0.00, spring accuracy = 0.93, fall RT = 0.00, spring RT = 0.62). It follows that the average leave-out mean of classmates' EF scores also increased between the fall and spring (fall class mean accuracy = 0.03, spring class mean accuracy = 0.94, fall class mean RT = 0.02, spring class mean RT = 0.62). Variability in classmates' EF scores decreased slightly for accuracy scores and increased slightly for RT scores during the school year (fall class *SD* accuracy = 0.87, spring class *SD* accuracy = 0.77, fall class *SD* RT = 0.94, spring class *SD* RT = 1.09). Students' processing speed also increased during the year (fall mean = 0.00, spring mean = 0.72). Due to multigrade classrooms, classroom average grade ranged from 2.58 to 5.00 ($M = 3.77$, $SD = 0.77$). The proportion of students eligible for FRPL varied across the schools in our study, ranging from 10% to 70.2% ($M = 0.43$, $SD = 0.23$).

Intraclass correlations (ICCs) test what proportion of the total variance in children's EF scores is accounted for by observed and unobserved factors operating at the classroom level. Therefore, a high ICC (close to 1.00, or 100%) would indicate high levels of selection bias, as children's EF scores would be entirely explained by the classrooms they are in. ICCs of the fall EF composite scores (accuracy ICC = 0.212, RT ICC = 0.088) demonstrate that a small proportion of the variance in students' EF scores at the beginning of the year

TABLE 1
Descriptive Statistics for All Study Variables

	Fall						Spring					
	<i>M</i>	<i>SD</i>	Min	Max	<i>n</i>	ICC	<i>M</i>	<i>SD</i>	Min	Max	<i>n</i>	ICC
Individual student EF measures												
Individual												
Accuracy	0.00	1.00	-3.44	2.33	722	0.212	0.93	0.83	-2.35	2.86	724	0.107
RT	0.00	1.00	-3.20	3.24	721	0.088	0.63	1.16	-3.56	3.67	724	0.098
Processing speed	0.00	1.00	-4.53	2.57	722	0.123	0.72	0.94	-2.87	2.88	724	0.107
Classroom-level EF measures												
Fall accuracy class												
Mean	0.03	0.49	-1.21	0.88	806		0.94	0.32	-0.02	1.62	806	
<i>SD</i>	0.87	0.20	0.39	1.40	806		0.77	0.18	0.31	1.15	806	
Fall RT class												
Mean	0.02	0.36	-0.80	0.83	806		0.62	0.43	-0.56	1.52	806	
<i>SD</i>	0.94	0.14	0.43	1.29	806		1.09	0.15	0.66	1.42	806	
Covariates												
Female	47.9%				802							
Age, months	9.899	0.828	8.000	12.37	806							
Class grade	3.770	0.769	2.580	5.000	33							
School FRPL	0.425	0.225	0.100	0.700	8							

Note. All EF scores are standardized per the fall scale. Individual EF composite ICCs are at the classroom level. Classroom means and standard deviations were calculated such that the individual student's score was not included in his or her score. RT scores were reversed such that higher scores can be interpreted as children demonstrating better EF skills (faster RTs). ICC = intraclass correlation; class = classmates'; RT = reaction time; FRPL = free or reduced-priced lunch (proportion of eligible students at the school).

was at the classroom level (e.g., between-classroom differences). This indicates that most of the variance in students' fall accuracy (78.8%) and RT (91.2%) scores was associated with individual child-level differences.

Bivariate correlations (Table 2) show that the leave-out mean and standard deviation were significantly and negatively correlated for accuracy in both the fall and the spring such that classrooms with higher mean accuracy scores had less variability in accuracy scores. Associations between classroom mean and classroom *SD* were not as highly correlated for RT scores. Individual students' processing speed was highly correlated with individual students' accuracy and even more highly correlated with individual students' RT. Older children and those in higher grades had higher EFs and less variability in EFs. Generally, schools with higher proportions of students receiving free or reduced-price lunch had lower EFs and more classroom variability in EFs.

Predicting Growth in EFs Through Measures of Classmates' EFs

As shown in Table 3, measures of classmates' accuracy on EF tasks were not associated with growth in accuracy scores (Model 1a; $\beta = 0.046, p = .688$). This was unchanged in Models 1b and 1c with the inclusion of individual students' fall RT and fall processing speed, respectively.

Individual students' fall accuracy scores were a strong predictor students' spring accuracy scores (Model 1a; $\beta = 0.507, p < .001$). Girls had significantly higher growth in accuracy scores across all models (Model 1a; $\beta = 0.154, p = .001$). Furthermore, students' fall RT scores had a trend-level association with growth in their accuracy scores ($\beta = -0.056, p = .057$, in Model 1b), suggesting that individual students' fall RTs uniquely contribute to their growth in accuracy scores over the year. Students' fall processing speed was not associated with growth in students' accuracy scores ($\beta = -0.032, p = .309$, in Model 1c). Variability in classmates' accuracy scores was not significantly associated with growth in students' accuracy scores and did not change results in any models.

However, RTs on EF tasks were associated with students' growth in speed on accurate trials from fall to spring (Table 3). It is important to note that RT scores have been reversed such that higher scores on the RT composite can be interpreted as better EFs. Model 2a shows that classmates' mean RTs were significantly associated with growth in individual students' RTs ($\beta = 0.332, p = .004$). This result remained significant with the inclusion of individual students' fall accuracy scores (Model 2b; $\beta = 0.305, p = .007$) and individual students' fall processing speed scores (Model 2c; $\beta = 0.300, p = .009$). Students whose peers had faster RTs, on average, responded more rapidly by the end of the school

TABLE 2
Bivariate Correlations Among All Study Variables

		Spring										
Fall		1	2	3	4	5	6	7	8	9	10	11
1	Acc individual	.61***	.27***	-.09*	.21***	.17***	-.02	.28***	.03	.19***	.22***	-.22***
2	Acc class mean	.41***	.78***	-.54***	.17***	.49***	-.21***	.22***	-.05	.49***	.57***	-.53***
3	Acc class SD	-.30***	-.72***	.50***	.10**	-.281***	.13***	-.14***	-.01	-.27***	-.26***	.06
4	RT individual	.31***	.17***	-.13***	.57***	.26***	-.04	.83***	-.23***	.27***	.25***	.04
5	RT class mean	.23***	.52***	-.34***	.24***	.85***	-.25***	.28***	-.04	.68***	.68***	-.10**
6	RT class SD	.09*	.16***	-.42***	.036	.15***	-.20***	-.09*	.00	-.23***	-.11**	.02
7	Processing speed	.52***	.25***	-.21***	.79***	.29***	.10**	.63***	-.23***	.28***	.26***	.05
8	Female	-.13***	-.06	.04	-.17***	-.02	.015	-.20**				
9	Age, months	.38***	.73***	-.50***	.28***	.61***	.16***	.31***	-.08*			
10	Class average grade	.40***	.81***	-.53***	.20***	.56***	.13***	.28***	-.07*	.76***		
11	School FRPL	-.25***	-.50***	.27***	.04	.11**	.11**	-.01	.06	-.19***	-.22***	

Note. The bottom section represents fall EF scores, and the top section represents spring EF scores. Shaded cells and bold font indicate correlations between variables in the fall and spring semesters. Classroom means and standard deviations were calculated such that the individual student's score was not included in his or her score. RT scores were reversed such that higher scores can be interpreted as children demonstrating better EF skills (faster RTs). Acc = accuracy; class = classmates; RT = reaction time; FRPL = free or reduced-priced lunch (proportion of eligible students at the school).
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE 3
Multilevel Models With Associations Between Fall and Classmates' EF Scores on EF Tasks With Spring EF Scores

	Accuracy			RT		
	Model 1a	Model 1b	Model 1c	Model 2a	Model 2b	Model 2c
Classroom mean EFs	0.046 (0.115)	0.057 (0.115)	0.053 (0.115)	0.332** (0.117)	0.302** (0.111)	0.300** (0.114)
Student fall						
Accuracy	0.507*** (0.030)	0.530*** (0.033)	0.523*** (0.034)		0.354*** (0.042)	
RT		-0.056+ (0.029)		0.607*** (0.038)	0.476*** (0.040)	0.333*** (0.058)
PS			-0.032 (0.031)			0.356*** (0.062)
Female student	0.154** (0.048)	0.143** (0.049)	0.145** (0.049)	-0.254*** (0.071)	-0.234*** (0.068)	-0.197** (0.071)
Average classroom grade	-0.056 (0.072)	-0.054 (0.072)	-0.055 (0.072)	0.091 (0.063)	-0.025 (0.061)	0.062 (0.062)
School-level FRPL	-0.198 (0.173)	-0.161 (0.174)	-0.172 (0.175)	-0.159 (0.168)	0.154 (0.166)	-0.204 (0.166)
Constant	1.188*** (0.266)	1.167*** (0.266)	1.176*** (0.267)	0.528+ (0.277)	0.820** (0.265)	0.634* (0.272)

Note. $N = 806$ students in 33 classrooms. All models presented account for the clustering of students within classrooms based on two-level multilevel regression analyses. RT scores were reversed such that higher scores can be interpreted as children demonstrating better EF skills (faster RTs). Model 1 predicts children's spring accuracy scores, and the classroom mean EFs represent classroom mean accuracy scores. Model 2 predicts children's spring RT scores, and the classroom mean EFs represent classroom mean RT scores. The inclusion of variability in classmates' EF scores (SD) did not change the results; therefore, these analyses are not presented. Values are presented as betas and standard errors in parentheses. EF = executive function; RT = reaction time; PS = processing speed; FRPL = free or reduced-priced lunch (proportion of eligible students at the school).
+ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

year as compared with students in classrooms whose peers had slower RTs.

Measures of individual students' RTs on EF tasks in the fall were significantly associated with their RTs in the spring across all models (Model 2a; $\beta = 0.607, p < .001$). In contrast with accuracy scores, boys demonstrated more growth in RT scores as compared with girls across all models (Model 2a; $\beta = -0.254, p < .001$). Students' fall accuracy scores were positively associated with growth in students' RT scores ($\beta = 0.354, p < .001$, in Model 1b). This shows that students who were more accurate on EF tasks in the fall responded more quickly by the end of the school year on the EF tasks. Furthermore, students' fall processing speed scores were positively associated with growth in students' RT scores ($\beta = 0.356, p < .001$, in Model 1c). This shows that both RT, representing children's EF abilities, and processing speed uniquely contribute to growth in RT scores. Variability in classmates' RT scores was not significantly associated with growth in students' RT scores and did not change results in any models.

Discussion

During the elementary school years, EFs play a key role in supporting children's engagement and adaptive behaviors in the classroom, as well as in learning academic skills (Finch & Obradović, 2017; Jacob & Parkinson, 2015; Nelson et al., 2017). Given the increasing importance and frequency of peer interactions in middle childhood (Gifford-Smith & Brownell, 2003), classmates' EFs may be a critical aspect of children's school experiences that influence children's self-regulation during this developmental period. This study provides initial evidence for an association between classmates' EFs and individual students' EFs among third, fourth, and fifth graders. Higher average levels of classmates' EFs, as measured by students' RTs on direct EF assessments, were linked to increased growth among individual students' RTs on EF tasks. In contrast, classmates' accuracy on the EF tasks was not associated with individual students' growth in accuracy scores.

Classmates' Accuracy Versus RT on EF Tasks

Our study revealed that students whose classmates had faster responses on direct assessments of EF skills, on average, demonstrated increased improvement in their speed over the academic year. These findings remained with the inclusion of children's processing speed, suggesting that these results are driven by EF-related speed on the tasks. Classmates' accuracy on EF tasks, however, was not associated with individual students' EF growth. There are several reasons why there may be discrepancies between these two metrics of EFs for children's development. First, it is possible that RT data are more sensitive to short-term changes in

children's EFs. Accuracy on the incongruent blocks of the MSIT and HF tasks showed indications of ceiling effects such that 19% and 34% of students attained perfect scores, respectively. Second, accuracy data are coarser than RT scores. When one is trying to capture small improvements in children's skills, it may be best to use a continuous measure of children's abilities, such as RTs, rather than ordinal accuracy scores. Therefore, RT scores may be able to more precisely capture improvements in EF skills over an 8-month period and may be particularly important for assessment of children in middle childhood. Alternatively, accuracy and RT scores on the same EF tasks may measure conceptually different aspects of EFs. Results demonstrated that students' fall RT scores contributed uniquely to growth in accuracy scores (at a trend level) and that students' fall accuracy scores contributed uniquely to growth in RT scores. These independent effects suggest that these measures capture unique aspects of children's EFs. Our results also highlight significant gender differences on the EF scores such that girls demonstrate more growth in accuracy scores and boys demonstrate more growth in RT scores. Additional research is needed to understand whether measures of accuracy and RT differentially affect social dynamics and classroom behaviors for children.

Classmates' EFs and EF Growth

Our results are consistent with the broader literature, which has demonstrated the importance of peers for elementary school students' academic and behavioral development (Gottfried, 2014; Hanushek et al., 2003; Thomas et al., 2011; Yudron et al., 2014). In addition, our findings corroborate an early childhood study demonstrating that classmates' average levels of EFs were associated with growth in individual preschoolers' EFs in a sample of children drawn from community-based school programs without a specific intervention component. The one study that did not find peer effects of EFs was conducted with low-income children involved in a preschool intervention program that significantly increased classroom quality and children's EFs (Weiland & Yoshikawa, 2013, 2014).

We did not find effects of variability in classmates' EFs for individual students' EF growth. This aligns with the two studies that examined variability in classmates' academic skills, which also found null effects (Burke & Sass, 2008; Hanushek et al., 2003). The one study that examined variability in classmates' mental health symptoms found negative effects of heterogeneity in classmates' externalizing behavior problems for individual children's behavioral development (Yudron et al., 2014). However, they used teacher-reported measures of children's behaviors, which tend to be highly skewed such that most students are rated as having no behavioral issues and then a small number of students are identified as displaying behavioral

problems (Miner & Clarke-Stewart, 2008; Sulik, Blair, Greenberg, & the Family Life Project Investigators, 2017). Therefore, classrooms with more variability in students' externalizing problems likely indicate that those classrooms have at least one student with externalizing behavioral problems. It is not surprising that this is negatively associated with individual students' behavioral development, as externalizing behavior problems are highly disruptive in the classroom context (Thomas, Bierman, & the Conduct Problems Prevention Research Group, 2006). In contrast, variability in EFs, as measured on direct assessments, would not likely have a strong impact on the classroom climate, given that these measures are normally distributed. Therefore, higher variability in classmates' EFs does not stand as a proxy for classrooms with children who have very low EFs.

Possible Mechanisms for Understanding Classmates' EFs

This study is the first to highlight that classroom measures of EFs play a role in children's EF development in middle childhood. Although we were not able to directly test the underlying mechanisms for our results, it is likely that our findings reflect processes among students and between students and teachers. Since the 1970s, researchers have highlighted that classroom environments should be viewed as social ecologies (Doyle, 1977; Vygotsky, 1978), where learning is a socially organized process (Kindermann & Vollet, 2014). Peer interactions and teacher-child interactions together shape children's opportunities to build and practice their self-regulation skills. Accordingly, we propose three mutually informative mechanisms that are all fruitful areas for future research.

Classmates' EFs and peer interactions. It is possible that being in a classroom with well-regulated peers may provide individual students with more examples of appropriate behavior and increased opportunities to engage in play and academic work that require self-regulation skills. Research focused on the development of children's behavioral problems demonstrates that children who are in classrooms with low average levels of aggression are less likely to establish friendships with aggressive peers, decreasing their own aggressive behaviors over the school year (Powers, Bierman, & the Conduct Problems Prevention Research Group, 2013). Children who are in classrooms with higher average levels of EFs may be more likely to interact and establish friendships with well-regulated peers, who motivate and support them to improve their own EFs.

Teachers' influences on peer interactions. Links between classmates' EFs and individual students' EFs may be also

mediated through teacher-student interactions (Farmer, McAuliffe Lines, & Hamm, 2011; Hughes & Chen, 2011). Teachers' behaviors and expectations shape peer interactions, as teachers establish classroom norms for what is considered acceptable and they model social interactions through their own relationships with students (Farmer et al., 2011). Supportive relationships between teachers and students have been linked to more positive interactional patterns among peers (Hendrickx, Mainhard, Boor-Klip, Cillessen, & Brekelmans, 2016; Hughes, Cavell, & Willson, 2001; Mikami & Mercer, 2017). It is plausible that individual relationships between teachers and students shape peer interactions in ways that increase or reduce opportunities for the development of EFs.

Students' EFs eliciting teachers' behaviors. Our results may also be explained by how children's behaviors influence teachers' instructional choices (Doyle & Ponder, 1975). Students with higher EFs demonstrate more learning-related behaviors, such as increased on-task behavior, more leadership skills, better prosocial skills, and decreased behavior problems (Finch & Obradović, 2017; Nelson et al., 2017; Sasser et al., 2015). Higher average levels of students' EFs may induce teachers to design activities that further benefit students' EF development. There is some evidence that elementary school students' self-regulation skills are linked to the quality of their interactions with teachers (Hernández et al., 2017; Portilla, Ballard, Adler, Boyce, & Obradović, 2014), but limited work has examined how students' self-regulation skills affect teachers' choices about learning activities in the classroom.

Limitations and Future Directions

Although our study represents an important effort to understand the role of classmates' EFs for children's developmental trajectories, it has several limitations that need to be addressed in future work.

First, our study does not examine the role of teachers' behaviors and teacher-student relationships in concert with classmates' EFs. Given studies demonstrating bidirectional associations between teachers' behaviors and peer group interactions (Farmer et al., 2011; Hughes & Chen, 2011; Luckner & Pianta, 2011), future work should build on our study to better understand how classroom measures of EFs affect teachers' instructional choices and behaviors.

Second, it is possible that our findings are driven by teachers' behaviors very early in the school year. Teachers who foster classroom climates that encourage good behavior and regulation may have influenced students' EF scores in the fall and growth in EFs over the school year. This seems unlikely, however, given that the fall assessments were completed within the first 3 weeks of the school year.

Third, it is possible that ceiling effects limited the sensitivity of our accuracy measures and explained the null findings for our accuracy composite. Longer and more challenging EF tasks might have yielded a similar pattern of results for both accuracy and RT measures. Future studies exploring peer effects in middle childhood should include more demanding tasks that produce more variability in children's accuracy scores.

Fourth, our study is correlational in nature and does not provide causal estimates of peer effects on students' EF development. It is possible that selection effects of students into classrooms and schools influence our findings. Within schools, students are nonrandomly sorted into classrooms. For example, teachers with higher scores on licensure exams tend to get fewer minority students; more affluent students, whose parents are more likely to be college graduates; and students who have higher average test scores (Clotfelter, Ladd, & Vigdor, 2006). Therefore, peer EFs may be a proxy for teacher quality such that students whose peers have higher EFs also have higher-quality teachers. To better explore the causal role of peers, future studies could use experimental designs to randomly assign students into teachers' classrooms with differing peer groups or leverage quasi-experimental designs.

Finally, our EF composite measures had relatively low reliability, likely because they are based on only three (RT) and four (accuracy) task-level variables. In future studies, including more EF tasks would help improve reliability on composite EF measures.

Conclusion

In summary, our findings suggest that classmates' EFs play a role in individual children's EF development during middle childhood. We highlight the need to identify specific aspects of classrooms that influence EF development in elementary school. Future studies should examine the dynamic interplay among classmates' EFs, teachers' behaviors, and interactions among teachers and students in the classroom to better understand the mechanisms explaining peer EF effects. This work demonstrating the importance of peer effects for EF development has implications for how students are assigned to classrooms and how teachers group students for learning activities.

APPENDIX TABLE A1

Skewness and Excess Kurtosis Values for RT Variables

Variable	Skewness	Excess kurtosis
Fall HF		
Incongruent RT	-0.05	-0.54
Mixed RT	-0.91	1.02

(continued)

APPENDIX TABLE A1. (continued)

Variable	Skewness	Excess kurtosis
Spring HF		
Incongruent RT	0.97	1.53
Mixed RT	0.02	-0.12
Incongruent RT		
Fall MSIT	-0.29	0.10
Spring MSIT	0.32	-0.03
Fall Flanker		
Blue RT	0.77	0.29
Pink RT	0.59	-0.20
Mixed RT	-0.61	0.32
Spring Flanker		
Blue RT	1.17	1.68
Pink RT	0.91	0.23
Mixed RT	-0.68	0.63

Note. HF = Hearts and Flowers; RT = reaction time; MSIT = Multi-Source Interference Test.

Authors' Note

This research was supported by a William T. Grant Foundation grant to Jelena Obradović. The preparation of this manuscript was also supported by a William R. and Sara Hart Kimball Stanford Graduate Fellowship to Jenna E. Finch.

Acknowledgments

We thank the children, teachers, and school administrators who participated and made this research possible and the students who helped collect and process the data.

ORCID iD

Jenna E. Finch  <https://orcid.org/0000-0003-0709-2318>

References

- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641–1660. doi:10.1111/j.1467-8624.2010.01499.x
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29*(3), 180–200. doi:10.1016/j.dr.2009.05.002
- Bierman, K. L., Torres, M. M., Domitrovich, C. E., Welsh, J. A., & Gest, S. D. (2009). Behavioral and cognitive readiness for school: Cross-domain associations for children attending Head Start. *Social Development, 18*(2), 305–323. doi:10.1111/j.1467-9507.2008.00490.x
- Blair, C., & Razza, R. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647–663. doi:10.1111/j.1467-8624.2007.01019.x
- Boucher, V., Bramoullé, Y., Djebbari, H., & Fortin, B. (2014). Do peers affect student achievement? Evidence from Canada using group size variation. *Journal of Applied Econometrics, 29*(1), 91–109. doi:10.1002/jae.2299

- Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematic achievement. *Developmental Science, 14*(4), 679–692. doi:10.1111/j.1467-7687.2010.01012.x
- Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics, 31*(1), 51–82. doi:10.1086/666653
- Bush, G., & Shin, L. M. (2006). The Multi-Source Interference Task: An fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network. *Nature Protocols, 1*(1), 308–313. doi:10.1038/nprot.2006.48
- Cartwright, K. B. (2012). Insights from cognitive neuroscience: The importance of executive function for early reading development and education. *Early Education & Development, 23*(1), 24–36. doi:10.1080/10409289.2011.615025
- Ciairano, S., Visu-Petra, L., & Settanni, M. (2007). Executive inhibitory control and cooperative behavior during early school years: A follow-up study. *Journal of Abnormal Child Psychology, 35*(3), 335–345. doi:10.1007/s10802-006-9094-z
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources, 41*(4), 778–820. doi:10.3368/jhr.XLI.4.778
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. doi:10.1016/j.neuropsychologia.2006.02.006
- Day, S. L., Connor, C. M., & McClelland, M. M. (2015). Children's behavioral regulation and literacy: The impact of the first grade classroom environment. *Journal of School Psychology, 53*(5), 409–428. doi:10.1016/j.jsp.2015.07.004
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135–168. doi:10.1146/annurev-psych-113011-143750
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*(5855), 1387–1388. doi:10.1126/science.1151148
- Doyle, W. (1977). Paradigms for research on teacher effectiveness. *Review of Research in Education, 5*(1), 163–198. doi:10.3102/0091732X005001163
- Doyle, W., & Ponder, G. A. (1975). Classroom ecology: Some concerns about a neglected dimension of research on teaching. *Contemporary Education, 46*(3), 183–188.
- Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy, 98*, 4–18. doi:10.1016/j.brat.2016.11.008
- Farmer, T. W., McAuliffe Lines, M., & Hamm, J. V. (2011). Revealing the invisible hand: The role of teachers in children's peer experiences. *Journal of Applied Developmental Psychology, 32*(5), 247–256. doi:10.1016/j.appdev.2011.04.006
- Figlio, D. (2005). *Boys named Sue: Disruptive children and their peers* (Working Paper No. 11277). Cambridge, MA: National Bureau of Economic Research.
- Finch, J. E., & Obradović, J. (2017). Independent and compensatory contributions of executive functions and challenge preference for students' adaptive classroom behaviors. *Learning and Individual Differences, 55*, 183–192. doi:10.1016/j.lindif.2017.03.002
- Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (2nd ed.). Hoboken, NJ: Wiley.
- Gifford-Smith, M. E., & Brownell, C. A. (2003). Childhood peer relationships: Social acceptance, friendships, and peer networks. *Journal of School Psychology, 41*(4), 235–284. doi:10.1016/S0022-4405(03)00048-7
- Gottfried, M. A. (2014). Peer effects in urban schools: Assessing the impact of classroom composition on student achievement. *Educational Policy, 28*(5), 607–647. doi:10.1177/0895904812467082
- Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*(1), 101–131. doi:10.1037/0033-295X.113.1.101
- Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527–544. doi:10.1002/jae.741
- Hendrickx, M. M., Mainhard, M. T., Boor-Klip, H. J., Cillessen, A. H., & Brekelmans, M. (2016). Social dynamics in the classroom: Teacher support and conflict and the peer ecology. *Teaching and Teacher Education, 53*, 30–40. doi:10.1016/j.tate.2015.10.004
- Hernández, M. M., Valiente, C., Eisenberg, N., Berger, R. H., Spinrad, T. L., VanSchyndel, S. K., . . . Thompson, M. S. (2017). Elementary students' effortful control and academic achievement: The mediating role of teacher-student relationship quality. *Early Childhood Research Quarterly, 40*, 98–109. doi:10.1016/j.ecresq.2016.10.004
- Hollenstein, T. (2007). State space grids: Analyzing dynamics across development. *International Journal of Behavioral Development, 31*(4), 384–396. doi:10.1177/0165025407077765
- Hollenstein, T., Granic, I., Stoolmiller, M., & Snyder, J. (2004). Rigidity in parent-child interactions and the development of externalizing and internalizing behavior in early childhood. *Journal of Abnormal Child Psychology, 32*(6), 595–607. doi:10.1023/B:JACP.0000047209.37650.41
- Holmes, C. J., Kim-Spoon, J., & Deater-Deckard, K. (2015). Linking executive function and peer problems from early childhood through middle adolescence. *Journal of Abnormal Child Psychology, 44*(1), 31–42. doi:10.1007/s10802-015-0044-5
- Hughes, J. N., Cavell, T. A., & Willson, V. (2001). Further support for the developmental significance of the quality of the teacher-student relationship. *Journal of School Psychology, 39*(4), 289–301.
- Hughes, J. N., & Chen, Q. (2011). Reciprocal effects of student-teacher and student-peer relatedness: Effects on academic self efficacy. *Journal of Applied Developmental Psychology, 32*(5), 278–287. doi:10.1016/j.appdev.2010.03.005
- Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research, 85*(4), 512–552. doi:10.3102/0034654314561338
- Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates' skills? *Child Development, 82*(6), 1768–1777. doi:10.1111/j.1467-8624.2011.01665.x
- Kindermann, T. A., & Vollet, J. W. (2014). Social networks within classroom ecologies: Peer effects on students' engagement in

- the context of relationships with teachers and parents. *Zeitschrift Für Erziehungswissenschaft*, 17(S5), 135–151. doi:10.1007/s11618-014-0555-9
- Kolkman, M. E., Hoijsink, H. J. A., Kroesbergen, E. H., & Leseman, P. P. M. (2013). The role of executive functions in numerical magnitude skills. *Learning and Individual Differences*, 24, 145–151. doi:10.1016/j.lindif.2013.01.004
- Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, 84(6), 1933–1953. doi:10.1111/cdev.12096
- Luckner, A. E., & Pianta, R. C. (2011). Teacher-student interactions in fifth grade classrooms: Relations with children's peer behavior. *Journal of Applied Developmental Psychology*, 32(5), 257–266. doi:10.1016/j.appdev.2011.02.010
- Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Development*, 80(3), 686–702. doi:10.1111/j.1467-8624.2009.01291.x
- Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, 101(3), 689–704. doi:10.1037/a0014240
- McKown, C. (2017). Social-emotional assessment, performance, and standards. *The Future of Children*, 27(1), 157–178.
- Mercer, S. H., McMillen, J. S., & DeRosier, M. E. (2009). Predicting change in children's aggression and victimization using classroom-level descriptive norms of aggression and prosocial behavior. *Journal of School Psychology*, 47(4), 267–289. doi:10.1016/j.jsp.2009.04.001
- Mikami, A. Y., & Mercer, S. H. (2017). Teacher behaviors toward children with attention-deficit/hyperactivity disorder predict peers' initial liking and disliking impressions in a summer camp setting. *Journal of Social and Clinical Psychology*, 36(6), 506–534. doi:10.1521/jscp.2017.36.6.506
- Miner, J. L., & Clarke-Stewart, K. A. (2008). Trajectories of externalizing behavior from age 2 to age 9: Relations with gender, temperament, ethnicity, parenting, and rater. *Developmental Psychology*, 44(3), 771–786. doi:10.1037/0012-1649.44.3.771
- Montroy, J. J., Bowles, R. P., & Skibbe, L. E. (2016). The effect of peers' self-regulation on preschooler's self-regulation and literacy growth. *Journal of Applied Developmental Psychology*, 46, 73–83. doi:10.1016/j.appdev.2016.09.001
- Muller, C. M., Hofmann, V., Fleischli, J., & Studer, F. (2016). Effects of classroom composition on the development of antisocial behavior in lower secondary school. *Journal of Research on Adolescence*, 26(2), 345–359. doi:10.1111/jora.12195
- Muthén, L. K., & Muthén, B. O. (2014). *Mplus (Version 7.3)*. Los Angeles, CA: Author.
- Nelson, T. D., Nelson, J. M., James, T. D., Clark, C. A. C., Kidwell, K. M., & Espy, K. A. (2017). Executive control goes to school: Implications of preschool executive performance for observed elementary classroom learning engagement. *Developmental Psychology*, 53(5), 836–844. doi:10.1037/dev0000296
- Obradović, J., Sulik, M. J., Finch, J. E., & Tirado-Strayer, N. (2018). Assessing students' executive functions in the classroom: Validating a scalable group-based procedure. *Journal of Applied Developmental Psychology*, 55, 4–13. doi:10.1016/j.appdev.2017.03.003
- Pedersen, S., Vitaro, F., Barker, E. D., & Borge, A. I. H. (2007). The timing of middle-childhood peer rejection and friendship: Linking early behavior to early-adolescent adjustment. *Child Development*, 78(4), 1037–1051. doi:10.1111/j.1467-8624.2007.01051.x
- Portilla, X. A., Ballard, P. J., Adler, N. E., Boyce, W. T., & Obradović, J. (2014). An integrative view of school functioning: Transactions between self-regulation, school engagement, and teacher-child relationship quality. *Child Development*, 85(5), 1915–1931. doi:10.1111/cdev.12259
- Powers, C. J., Bierman, K. L., & the Conduct Problems Prevention Research Group. (2013). The multifaceted impact of peer relations on aggressive/disruptive behavior in early elementary school. *Developmental Psychology*, 49(6), 1174–1186. doi:10.1037/a0028400
- Sasser, T. R., Bierman, K. L., & Heinrichs, B. (2015). Executive functioning and school adjustment: The mediational role of pre-kindergarten learning-related behaviors. *Early Childhood Research Quarterly*, 30, 70–79. doi:10.1016/j.ecresq.2014.09.001
- Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology*, 41(3), 457–471. doi:10.1007/s10802-012-9684-x
- Skibbe, L. E., Phillips, B. M., Day, S. L., Brophy-Herb, H. E., & Connor, C. M. (2012). Children's early literacy growth in relation to classmates' self-regulation. *Journal of Educational Psychology*, 104(3), 541–553. doi:10.1037/a0029153
- Sulik, M. J., Blair, C., Greenberg, M., & the Family Life Project Investigators. (2017). Child conduct problems across home and school contexts: A person-centered approach. *Journal of Psychopathology and Behavioral Assessment*, 39(1), 46–57. doi:10.1007/s10862-016-9564-8
- Thomas, D. E., Bierman, K. L., & Powers, C. J. (2011). The influence of classroom aggression and classroom climate on aggressive-disruptive behavior. *Child Development*, 82(3), 751–757. doi:10.1111/j.1467-8624.2011.01586.x
- Thomas, D. E., Bierman, K. L., & the Conduct Problems Prevention Research Group. (2006). The impact of classroom aggression on the development of aggressive behavior problems in children. *Development and Psychopathology*, 18(2), 471–487. doi:10.1017/S0954579406060251
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., . . . Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2–3), 119–145. doi:10.1177/016235320302700203
- Véronneau, M.-H., Vitaro, F., Brendgen, M., Dishion, T. J., & Tremblay, R. E. (2010). Transactional analysis of the reciprocal links between peer experiences and academic achievement from middle childhood to early adolescence. *Developmental Psychology*, 46(4), 773–790. doi:10.1037/a0019816
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. *Early Childhood Research Quarterly*, 28(2), 199–209. doi:10.1016/j.ecresq.2012.12.002

- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84*(6), 2112–2130. doi:10.1111/cdev.12099
- Weiland, C., & Yoshikawa, H. (2014). Does higher peer socio-economic status predict children's language and executive function skills gains in prekindergarten? *Journal of Applied Developmental Psychology, 35*(5), 422–432. doi:10.1016/j.appdev.2014.07.001
- Williford, A. P., Vick Whittaker, J. E., Vitiello, V. E., & Downer, J. T. (2013). Children's engagement within the preschool classroom and their development of self-regulation. *Early Education & Development, 24*(2), 162–187. doi:10.1080/10409289.2011.628270
- Yudron, M., Jones, S. M., & Raver, C. C. (2014). Implications of different methods for specifying classroom composition of externalizing behavior and its relationship to social-emotional outcomes. *Early Childhood Research Quarterly, 29*(4), 682–691. doi:10.1016/j.ecresq.2014.07.007
- Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education* (No. NCER 2017-2000). Washington, DC: National Center for Education Research.

Authors

JENNA E. FINCH is an assistant professor of psychology at the University of Nebraska–Lincoln. Her research interests include executive functions, motivation, early childhood education, and contextual influences on children's cognitive and social-emotional development.

ELISA B. GARCIA is an early childhood researcher in the education division at SRI International. Her research addresses early childhood development, with a focus on dual language learners and early education.

MICHAEL J. SULIK is a postdoctoral scholar in the Graduate School of Education at Stanford University. He studies the development of children's self-regulation.

JELENA OBRADOVIĆ is an associate professor in the developmental and psychological sciences program in the Graduate School of Education at Stanford University. Her research examines how the interplay of children's physiological arousal, self-regulatory skills, and the quality of caregiving environments contributes to children's adaptation over time.