Small-Group Composition Effects on Executive Function in Early Elementary School

Stuart Hannah

Department of Psychology, The University of Michigan

PSYCH 426: Senior Honors Research II for Psychology as a Natural Science

Dr. Frederick Morrison

April 12, 2020

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science with Honors in Psychology from the University of Michigan 2020

Author Note

This research was supported by the National Science Foundation (Grant #1356118).

Correspondence concerning this article should be addressed to Stuart Hannah, The Pathways to Literacy Lab, University of Michigan, 3242 East Hall, 530 Church St, Ann Arbor, MI 48109. Email: stuahann@umich.edu

Abstract

Executive function (EF) skills have been found to be particularly important for regulating behavior and subsequent math and literacy achievement in elementary school. Understanding classroom environmental factors that maximize EF skills is crucial for future educational interventions. Since small group activities are so common in elementary school classrooms, an environmental factor worth examining is the composition of these groups, mainly two characteristics, size and gender ratio. In this study, we examined the effects of group size and gender ratio on group-based measures of EF skills in kindergarteners (N = 251). We did not find a significant main effect of group size on any of the group-based EF measures, even after controlling for the relevant covariates. Our results showed no significant main effect of gender ratio on any of the group-based EF measures, and no significant correlations between gender ratio and performance on any of these measures. However, we did find significant main effects of gender on the group-based measures of attention and working memory. Future research is necessary to confirm the results of this study. We hope that these findings will help inform small-group selection in kindergarten, allowing teachers more flexibility in creating groups without concern for the size or gender ratio of each group.

Keywords: executive function, self-regulation, group composition, peer effects, group size, gender differences

Small-Group Composition Effects on Executive Function in Early Elementary School

Many of the obstacles teachers face instructing children in early elementary school stem from their students' lack of self-regulation skills (Zelazo et al., 2016). When young students are moderately attentive, undistracted, and sitting calmly in their seats, learning flourishes. Although these conditions are crucial for learning, they are far from guaranteed in early elementary school classrooms, often requiring every ounce of a teacher's training and willpower to achieve. Understanding this predicament, educational researchers over the years have investigated many of the neurocognitive processes necessary for young children to regulate their behavior.

The set of neurocognitive processes generally considered to be responsible for these abilities are executive functions (EFs; Zelazo et al., 2016). Although many different processes have been studied and linked to EF, the three main components of EF are understood to be inhibitory control, attention, and working memory (McClelland et al., 2007a; Morrison et al., 2010). It is useful to measure each component of EF separately since they are distinct functions related strongly to different academic skills within math and reading (Blair & Razza, 2007). However, they are moderately correlated to each other and may be utilized at the same time. Thus, EFs are considered to be part of a unified construct (Miyake et al., 2000).

Although the importance of EF for academic achievement in early elementary school has been well established by prior research (McClelland et al., 2006; McClelland et al., 2007a; Morrison et al., 2010; Zelazo et al., 2016), much of our understanding relies on lab-based measures of EF (McCoy, 2019). While lab-based measures of EF have been essential in determining a link between EF and academic achievement, they cannot be used to inform researchers on group context factors that may affect the manifestation of EF skills in the classroom. Thus, reliance on lab-based measures by prior research has led to a gap in the research literature concerning the effects of group context factors on EF. This information is particularly relevant when considering educational interventions in early elementary which target children's self-regulation skills.

Recently, developmental researchers have provided new group-based measures of EF which can capture the effects of group context factors while allowing a sufficient amount of experimental control to study them (Ahmed et al., 2018; McCabe & Brooks-Gunn, 2007; Obradović et al., 2017). Since these measures were recently made available, little research has been conducted using them to study specific group context factors and determine their effects on EF. One aspect of the group which has yet to be studied with respect to EF in early elementary-aged children is group composition. Considering the widespread use of small-group activities in early elementary classrooms, understanding the effects of group composition on students' EF skills may be able to inform how teachers select the groups. The aim of this study was to use new group-based EF measures to investigate the effects of group size and gender ratio on EF skills in kindergarteners.

EF's Impact on Academic and Life Outcomes

Prior research has demonstrated a notable link between academic achievement and behavioral self-regulation skills, beginning in preschool and persisting throughout all formal schooling (McClelland et al., 2006; McClelland et al., 2007a; Morrison et al., 2010; Zelazo et al., 2016). Measured EF skills have been shown to strongly predict math and literacy skills in preschool, kindergarten, and throughout elementary school (McClelland et al., 2007b; Blair & Razza, 2007; McClelland et al., 2006). Additionally, the gap between children with weak self-regulation skills and children with strong self-regulation skills expands significantly between kindergarten and 2nd grade, which suggests that early elementary school years are a critical time for EF development (McClelland et al., 2006).

Self-regulation skill deficiencies which are not addressed during early childhood can persist throughout schooling and contribute to a range of negative outcomes including: early smoking, dropping out of school, teen pregnancy, sexually transmitted disease, substance abuse, criminal behavior, and financial struggles (Moffitt et al., 2011). In response to these and other findings indicating the importance of behavioral self-regulation in educational and life outcomes, researchers from a wide range of fields, such as developmental psychology, neuroscience, education, and public policy, continue to study EF. Particular attention in these fields has been drawn to identifying environmental influences on EF and interventions designed to improve EF skills in a variety populations (Zelazo et al., 2016).

Akin to other areas of cognitive development, early intervention is critical for resolving EF deficiencies since individual differences in EF skills like working memory emerge as early as age 3 (Wiebe et al., 2011), predict math and reading proficiency until the age of 15 (Ahmed et al., 2018; Watts et al., 2014), and predict health, wealth, and public safety at the age of 32 (Moffitt et al., 2011). Some educational researchers consider early behavioral self-regulation interventions to be an option in working towards closing the achievement gap in education. This coming after research demonstrating that racial minority children and those with low socio-economic status (SES) tend to score lower on measures of EF. These differences can explain the racial and SES gaps in early academic achievement (Nesbitt et al., 2013).

EF Measures

Although overall classroom quality, measured by the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2007), has been strongly linked to EF skills in multiple studies

(Weiland et al., 2012; Rimm-Kaufman et al., 2009), few if any studies have been able to identify which specific classroom environmental variables enhance EF skills in early elementary students. Until recently, strongly validated EF measures were not sufficient for this kind of exploration, which may explain this gap in the literature. To this point, the majority of EF measures used in studies have been classroom observations, teacher reports, and lab-based, one-on-one measures. These do not allow for the experimental study of group context factors that influence EF. On the one hand, classroom observation allows for the naturalistic study of group context factors (a benefit not gained from individual lab-based measures). On the other, classroom observation does not allow experimental control and manipulation of the classroom factors, so it is difficult to determine the causality of factors affecting EF skills of students (McCoy, 2019). Ultimately, this leads to a lack of studying classroom variables experimentally with EF measures.

Although there is a dense research literature on the topic of EF and its role in early childhood academic achievement, it relies heavily on studies using individual lab-based assessments to measure EF. For instance, one of the most commonly used EF measures in research literature today is the Head-Toes-Knees-Shoulders task (HTKS) (Ponitz et al., 2008). HTKS has been validated by many studies for its strong link to multiple EF components, including cognitive flexibility, inhibitory control, and working memory (McClelland et al., 2014). However, like many other lab-based measures, this measure is conducted in a one-on-one setting, rather than a classroom setting among peers. Although measures like HTKS may capture these EF components with superior experimental control, they do not capture EF skills in the context in which they are most often used, the classroom.

Upon recognizing this gap in the literature, researchers produced ecologically valid, experimentally controlled measures of EF in the form group-based assessments or games (Ahmed et al., 2018; McCabe & Brooks-Gunn, 2007; Obradović et al., 2017). Although these EF measures are still new and minimally tested, they allow researchers to measure EF in the classroom while still retaining experimental control over variables of study. This provides researchers with a measure somewhere in between classroom observation and lab-based measures, marrying some of the benefits of each approach. This is imperative for determining whether or not environmental factors play a substantial role in the expression of EF skills in the classroom. Initial research involving the use of group-based measures suggests that environmental factors do play a role.

Group-based measures in a classroom context reveal that environmental factors, like peer influences, have a strong impact on the manifestation of EF skills (McCabe & Brooks-Gunn, 2007). McCabe and Brooks-Gunn (2007) demonstrated the viability of using group-based measures to study self-regulation and allowed for the examination of peer influences on children’s EF skills. They produced multiple ecologically valid measures of EF for children in the form of the *Snack Delay*, *Gift Wrap*, and *Walk-a-Line* tasks. They used these EF measures to test young children’s delay of gratification in individual and group settings. They found that children showed significantly less self-regulatory control across all three tasks in the group context. In addition, girls performed significantly better than boys on the *Gift Wrap* task. Boys becoming more easily distracted by peers could explain why they consistently underperform with respect to girls on ecologically sensitive measures, such as CBRS and the *Gift Wrap* task, while still performing adequately on lab-based achievement measures (Duckworth & Seligman, 2006; Matthews et al., 2009; McCabe & Brooks-Gunn, 2007). Understanding peer effects on EF skills could be the key to successful EF intervention in the classroom. Although peer distractions in the classroom have been shown to be particularly detrimental to attention (Felmlee, Eder, & Tsui, 1985), researchers require a clear understanding of their effect on EF skills in order to produce empirically based interventions to minimize them.

Gender Differences in EF

In many studies investigating individual differences in EF skills and academic achievement, researchers have found significant gender differences (Duckworth & Seligman, 2006; Matthews et al., 2009). In these studies, researchers found that girls generally performed better than boys on EF and self-regulation tasks. For example, in a longitudinal study of 268 kindergarteners, girls performed significantly better than boys on multiple EF measures, and there was a significant positive relationship between EF and achievement measures (Matthews et al., 2009). Researchers used Child Behavior Rating Scale (CBRS) teacher reports and Head-Toes-Knees-Shoulders (HTKS) to measure self-regulatory skills, and the Woodcock–Johnson III Tests of Achievement to measure academic achievement (Bronson et al., 1995; Ponitz et al., 2008; Woodcock et al., 2001). However, girls did not score significantly higher than boys on the Woodcock-Johnson III Test of Achievement measures (Matthews et al., 2009). These results mirror the findings of another EF study, Duckworth and Seligman (2006), which found that girls performed significantly better than boys on measures of self-discipline but not on measures of achievement test scores. These findings could be explained by boys having a stronger sensitivity to environmental distractors in the classroom than girls. However, more research is needed to establish a link between specific environmental influences and weakened EF skills.

Small-Group Composition

A popular way teachers have sought to increase academic achievement is through dividing the classroom into smaller groups. Research has shown that grouping students for instruction provides a significant advantage when compared to whole-class instruction, and the minimization of peer effects mediates the relationship between group composition and students' learning (Wilkinson & Fung, 2002). Furthermore, researchers have suggested that the learning benefits of small-group instruction can be increased by altering the composition of the groups. One of the most common methods for instruction grouping in elementary school is within-class ability grouping (Chorzempa & Graham, 2006). Homogeneous ability grouping (placing students into groups with other students of similar ability) has been shown to produce some benefits for overall learning (Lou et al., 1996). However, this form of grouping is quite controversial with many critics arguing that these benefits may be at the expense of the students in low-ability groups. These students may fall further behind their peers, face social stigmatization, spend less time in instructional activities, and receive lower quality instruction (Chorzempa & Graham, 2006; Sorensen & Hallinan, 1986).

Given the struggles to find the right balance between maximizing overall learning and minimizing the achievement gap while using ability grouping, there may be room to explore other group characteristics which have the potential to further overall learning without harming low performing students. Rather than grouping students based on academic ability, it may be useful to group students based on other characteristics which maximize their capacity to regulate their behavior. In many ways, EF researchers are exceptionally equipped with new group-based measures to study these group characteristics, particularly in elementary school (Ahmed et al., 2018; McCabe & Brooks-Gunn, 2007; Obradović et al., 2017). Group-based measures allow for the direct study of elementary school students participating in small group activities as well as the experimental control required to determine causal links between group characteristics and EF skills. Based on previous educational research, two of the group characteristics worth examining are group size and gender ratio.

Group Size

When constructing groups for small-group activities, elementary school teachers of all levels choose groups of a large variety of sizes. In each activity, there may be groups as large as 7 to 10 children and groups as small as 2 to 3 children. Since there is not a clearly established best size for small groups, teachers may select groups of different sizes for any number of lessons, but these groups are often based on table seating arrangements. In a review of small group instruction in elementary school, about two-thirds of small group observations consisted of 4 to 6 children and about one-third consisted of 7 to 10 children. These results were relatively consistent throughout early, middle, and late years of elementary school (Kutnick & Blatchford, 2013). Although previous research describes the size of groups typically used in elementary school, the relationship between small-group size and measures of EF or academic achievement is still unknown.

However, the large body of research literature demonstrating the effects of class size on academic achievement, attentiveness, and off-task behavior may provide insight into the potential effects of group size on group-based EF measures. One of the largest, most cited studies on the effects of class size on academic achievement is the Tennessee Student Achievement Ratio (STAR) experiment (Mosteller, 1995). In this randomized experiment of about 6,500 students in 330 K-3 classrooms in Tennessee, researchers found that students in small classes (13 to 17 students) scored significantly higher on academic achievement measures than students in large classes (22 to 25 students). Other important studies have since supported the conclusion that smaller classes in elementary schools produce better academic outcomes (Schanzenbach, 2014). Furthermore, in a review of class size literature, Finn, Pannozzo, and Achilles (2003) linked the improvement in academic achievement in small classes to student behavior, citing an increase student academic engagement as the primary factor. A more recent study by Blatchford, Edmonds, and Martin (2003) pointed specifically to a significant increase in off-task behavior associated with large elementary classes. In the context of EF research from classroom observations, there is a significant link between low response inhibition and working memory skills and off-task behavior (Moffett & Morrison, 2019). These findings in conjunction with the extensive research documenting the link between EF and academic achievement (Ahmed et al., 2018) may suggest that larger group sizes decrease students' EF skills resulting in off-task behavior and lower academic performance.

Gender Ratio

Due to the lack of group-based measures of EF, the peer effects of gender within small-group activities or classroom instruction are largely unexplored. Nonetheless, knowledge of gender differences found in previous EF studies can shed light on what the peer effects of gender may look like. As discussed earlier, girls outperform boys on average across many EF measures including teacher reports, lab-based measures, and newly developed group-based EF measures (Duckworth & Seligman, 2006; Matthews et al., 2009; Ahmed et al., 2018; McCabe & Brooks-Gunn, 2007). In addition, the lack of a disparity in academic achievement measures paired with the gender differences in ecologically sensitive EF measures suggest that boys may be more affected by EF environmental factors. Assuming this is the case, groups with a higher proportion of boys may produce more off-task behavior and distractions.

Current Study

For purposes of this study, we used a set of classroom group-based measures recently developed to include three components of EF: response inhibition, attention, and working memory. These measures include The *Freeze Game*, the *Jumping Game*, and the *Marching Game* (Ahmed et al., 2018). Each of the three tasks targets a specific component of EF: response inhibition, attentional control, or working memory. This configuration will allow us to understand how each component of EF is affected by small-group composition.

Although the current study selected groups based on availability of students, the composition of the groups participating in these tasks varied by gender ratio and group size. This will allow us to investigate the relationships between these specific group characteristics and EF skills. Due to the lack of many group-based measures of EF, there has not been research done on the effect of group size on EF measures. However, much research has shown that larger class sizes increase off-task behavior and negatively affect academic achievement (Finn et al., 2003; Blatchford et al., 2003; Mosteller, 1995; Schanzenbach, 2014). Since there are strong links between academic achievement, off-task behavior and EF, we expect the size of the group to have a similar effect on measures of EF. For our first research question, we hypothesize that larger groups in EF tasks would produce stronger peer effects resulting in lower individual EF performances. We will also determine whether or not this effect depends on the gender of the participant. Since girls in early elementary generally perform better in self-regulation tasks (Duckworth & Seligman, 2006; Matthews et al., 2009), we predict that they will be more resilient to these small-group composition effects. For our second research question, we hypothesize that participants in groups with a larger proportion of boys will score lower across all of the group games EF measures. We will also determine whether or not this depends on the gender of the participants. Since we anticipate girls in the study will be more resilient to peer effects in general, we expect that their scores on EF tasks will not be affected as much as boys by the size of the group.

Method

Participants

This is part of a 4-year longitudinal study following kindergarteners through second grade adding a new cohort of kindergarteners each year. The current study focused specifically on data collected from all four cohorts of kindergarteners which amounted to 251 participants. The ages of the children range from 5.0 years old to 7.0 years old (*M* = 5.8, *SD* = 0.37). For purposes of the group-based measures, kindergarteners were divided up into groups of 3 to 7 students (*M* = 4.8, *SD* = 0.81). This is particularly relevant as we will consider effects of group composition as well as individual differences in EF measures. This sample was selected from seven elementary schools in Southeastern Michigan. The percentages of students eligible for free or reduced-priced lunch by school were 2%, 33%, 39.9%, 42.4%, 61%, 68.5%, and 71.9%. Students were recruited for the study through elementary schools by contacting their parents to determine if they would like for their child to participate. Parents and teachers of the students gave consent to allow them to participate in the study and to be recorded on video for research uses.

Materials/ Procedures

The lab-based and group-based measures of EF were facilitated by trained research assistants in classrooms of the schools where the students were enrolled. For the lab-based measures, research assistants asked permission from teachers to pull out each student from class to complete each of the tasks in a separate room. For the group-based measures, research assistants brought with them two cameras on stands, a stereo, and tape for the floor. When they arrived, they were directed to an available classroom where they set up the cameras, a stereo, and placed a circle of tape on the floor. A group of 3 to 7 students was chosen based on availability of each student and whether or not they had already participated. Once a group of students was chosen by the research assistants, they went to classrooms in session to ask teachers if they could pull out students who were participating. Students were asked if they would like to play the group games before they started. The research assistants taught students how to play the games, and the games lasted between 8 and 15 minutes from start to finish. After they finished, the students were brought back to their classes.

Measures

Lab-Based Measures

Response Inhibition. To measure response inhibition, we used the Head-Toes-Knees-Shoulders task (HTKS) (Ponitz et al., 2008). Although HTKS also taps working memory and attention (McClelland et al., 2014), its linked more strongly to response inhibition than the other two lab-based measures when working with this three-factor EF structure (Ahmed, 2019). This task is similar to the popular childhood game *Simon Says*, except they are asked to do the opposite of what the instructor says every time. For example, if the instructor says, "Touch your head," students are expected to touch their toes. If the instructor says, "Touch your shoulders," students are expected to touch their knees. After each round, students are given a score of 0 for an incorrect response, 1 for a self-correction, and 2 for a correct response. Students completed 30 trials, and the maximum possible score was 60. Inter-rater reliability for this measure was 100%, and the internal consistency (Cronbach's α) was 0.83.

Working Memory. The lab-based, individual measure for working memory used in this study was Backward Digit Span, adapted from the Wechsler Intelligence Scales for Children (Wechsler, 1991). Backward Digit Span has been validated as a measure for working memory since participants are required to store information in their short-term working memory to accomplish the task. Students are given a list of numbers by the instructor and are asked to repeat the numbers backwards. The list starts with one number and increases with each round until the maximum of six numbers. The research assistant implementing the test stops and records a score for a student once they have failed two times consecutively (Nesbitt et al., 2013). After two full sets, the students received scores for the longest string of numbers they could recite backwards.

Attentional Control. In order to measure students' attentional control, we used the Pair Cancellation test from the Woodcock-Johnson III Tests of Cognitive Abilities (Woodcock et al., 2001). In this task, students are given a sheet of paper with a list of pictures in rows. Each picture is a cup, a ball, or a dog. During this task, students are asked to circle every pair consisting of a ball followed by a dog as quickly as possible. There are 69 correct pairs in the list and their score is recorded as the total number of pairs they are able to circle in 3 minutes.

Group-Based Measures

Freeze Game. In the *Freeze Game*, a group of 3-7 children are asked to start walking in a circle marked by a line of tape, until they hear music playing at which point, they freeze where they were. Once the music starts playing again, the students are instructed to begin walking in a circle again. The research assistants controlled the music starting it after intervals of about 7-15 seconds. One trial was recorded as walking once and freezing once and this continued for three trials. The trials were videotaped, and research assistants measured the stop-time and distractibility of students using video editing software. Stop-time was measured from the moment the music started playing to the moment when the child's front foot is flat on the floor and they stop moving. Distractibility was coded on a Likert Scale from 0 to 3 with a score of 0 signifying no attention to the task and a score of 3 signifying full attention to the task. Inter-rater reliability for the *Freeze Game* was 94%.

Jumping Game. In the *Jumping Game*, the same group of children are told new instructions. They are asked to walk in circle again, except this time they are told to walk when the music is playing and stop when the music stops playing. Before the first trial, students are instructed to stop and jump once when the music stops playing. Before the second trial, students are instructed to jump twice when the music stops and so on until three trials are complete. Stop time and distractibility are measured the same way as the first game. However, for this game the extra action of jumping is coded for action recall and action performance. Action recall is defined as whether or not the student recalled to jump on their own, they were cued visually or verbally by a peer to jump, or they did not jump at all. This scale includes a score of 0 if the student does not jump and a score of 4 if the student does jump without a cue. Action performance is defined as how many times the child jumps over or under the expected amount. Thus, if they jumped three times when they were expected to jump twice, they received a score of 1. A perfect score on this measure was a 0. Inter-rater reliability for the *Jumping Game* was 92%.

Marching Game. In the *Marching Game*, the same group of students were given new instructions once again. In this task, students are given instructions for two different songs. Before the first trial, the research assistants instruct the students to walk around the circle clockwise when song one is playing and walk around the circle counterclockwise when song two is playing. When the music stops, they are instructed to stop and face the center of the circle. The marching game adds marching recall and marching performance to the measures from the previous games. Marching recall was scored as 0 or 1 based on whether or not the child was cued to march in the direction they walked. Marching performance was scored on a Likert scale from 0 to 2. The child received a 0 if they marched in the wrong direction, a 1 if they marched in the right direction after being cued, and a 2 if they marched in the correct direction on their own. Inter-rater reliability for the *Marching Game* was 86%.

Analytic Plan

Using data from the lab-based, individual measures and the group-based measures of EF, we assessed whether or not the students performed worse on EF tasks when placed in a larger group or when placed in a group with more boys than girls. First, we ran descriptive statistics for all of our variables of interest including gender, group size, proportion of boys, all of the lab-based measures, and all of the group-based measures of EF. The group-based measures were calculated by adding together scores across games to create overall composite scores. Thus, the stop-time composite was calculated by combining a student's stop times from the *Freeze Game*, the *Jumping Game*, and the *Marching Game*. The same procedure was followed for calculating the distractibility and working memory composites.

Next, we fitted multivariate regression models for each of our research questions in order to determine main effects, control for covariates, and investigate whether or not there were interactions by gender. In each of our models, we predicted performance on one of the group-based measures of EF (stop-time composite, distractibility composite, or working memory composite) while controlling for the covariates (age at group games testing, school attended, gender, and the corresponding lab-based measure of EF). We controlled for these covariates due to previous research demonstrating significant effects of each of these variables on measures of EF (McClelland et al., 2006; Matthews et al., 2009; Zelazo et al., 2016).

For our first research question on the effect of group size, we fitted two models for each group-based measure. The first model included the predictor of group size in addition to the covariates (age at group games testing, school attended, the analogous lab-based measure, and gender). The second model included group size and all of the same covariates with the addition of the interaction term (group size x gender). For our second research question on the effect of the proportion of boys, we ran the same sequence of regression models, only replacing the variable of group size with proportion of boys. All statistical analyses were performed using IBM SPSS Statistics version 24.

Results

Descriptive statistics for every variable of interest are shown in Table 1. Pearson's bivariate correlations between these variables are provided in Table 2. With regard to the first research question, group size was only significantly correlated with age at testing and the working memory composite (the jumping game action performance/ recall). Furthermore, the correlation between group size and the working memory composite ($r = 0.15$, $p = .021$) was positive with students in larger groups actually performing better on this task. This was not consistent with our hypothesis that there would be a significant negative relationship between group size and all of the group-based EF measures.

Considering the second research question, proportion of boys was not significantly related with performance on any of the group-based EF measures. This did not support our hypothesis that students in groups with a higher proportion of boys would perform worse on all of the group-based EF measures. Although group size and proportion of boys did not have significant correlations with the group-based EF measures, significant gender differences were found in the lab-based measure HTKS ($r = 0.21$, $p = .001$) and the group-based measures distractibility composite ($r = 0.24$, $p < .001$) and working memory composite ($r = 0.15$, $p = .021$).

Overall, the results from the multivariate regression models did not support our hypotheses that group size and proportion of boys would have significant negative effects on the group-based EF measures. Our results for the stepwise regression models are provided in Tables 3-8. For the first research question, we examined models including group size as a predictor of the group-based EF measures, as shown in Tables 3-5. We did not find a significant main effect of group size on any of the group-based EF measures. In looking for an interaction due to gender, we still did not find any significant main effects of group size on the group-based measures of EF. However, we did find significant main effects of gender on the distractibility composite ($\beta = 0.20$, $p = .002$) and on the working memory composite ($\beta = 0.18$, $p = .004$). These findings suggest that girls significantly outperformed boys on these two tasks. There was no significant main effect of gender on the stop-time composite. Girls still performed better than boys in this area but the difference wasn't significant ($r = .11$, $p = .08$). Despite the main effects of gender on the distractibility and working memory composites, there were no significant interactions between group size and gender for predicting performance on any of the group-based EF measures.

For the second research question, we considered the models including proportion of boys as a predictor of the group-based EF measures, as shown in Tables 6-8. We did not find a significant main effect of proportion of boys on any of the group-based EF measures. The only significant relationship with respect to this question was between proportion of boys and performance on Backward Digit Span. The participants selected in groups with a higher proportion of boys had slightly higher performance on Backward Digit Span. When controlling for all covariates in the regression analysis, it appears that this had little to no effect on the results. While determining whether or not there was an interaction between proportion of boys and gender, once again, we did find significant main effects of gender on the distractibility and working memory composites similar to those found in the first research question. However, there was not a significant interaction between proportion of boys and gender. Thus, the proportion of boys in the group and group size did not seem to affect one gender's group-based EF performance more than the other.

For both research questions, we compared the models with and without the corresponding lab-based measures as covariates. Across all group-based EF measures, the models including related lab-based measures explained more of the variance in group-based EF performance. This demonstrated that the lab-based measures explained some of the variance in each of the corresponding group-based measures. Thus, controlling for the lab-based measures appears to have been appropriate.

Discussion

The current study sought to use new group-based EF measures to explore the effects of small-group characteristics on response inhibition, attention, and working memory in a group context. As far as we know, this is the first study investigating the effects of group size and gender ratio on group-based EF measures. It is our expectation that these group-based EF measures will allow further exploration into peer effects on EF in a small group setting. Despite our anticipation of larger groups resulting in lower performance on group-based EF tasks, we did not find a significant effect of group size on any of these measures. In addition, there was no significant interaction between group size and gender for predicting EF performance. We also did not find a significant effect of gender ratio on any of the group-based EF measures or a significant interaction between gender ratio and EF performance. Thus, the effects of these small-group characteristics on group-based EF performance did not depend on the gender of the child.

Group Size and EF Performance

The first goal of this study was to examine the role of group size in children's EF performance. Although we anticipated that larger groups would result in lower scores across the group-based EF measures, group size was only significantly correlated with age at group games testing and the working memory composite. By chance, it appears that students selected for the larger groups tended to be significantly older on average. This is not likely to have influenced our results since group size was still not a significant predictor of the group-based measures when controlling for age at testing during the regression analysis. As for the significant positive relationship between group size and the working memory composite, we expected this relationship to be in the opposite direction. This finding may be explained by the way the *Jumping Game* is structured. In a larger group, there is a better chance for one student to understand the rules then jump and clap the correct number of times. Students who follow the leader can still receive 3 out of the 4 points for performance. This could result in stronger performance on the working memory composite for students in larger groups.

We did not find a significant main effect of group size on any of the group-based EF measures even after controlling for the covariates (age at group games testing, school attended, gender, and the corresponding lab-based measures of EF). Thus, it appears from our findings that kindergarten students’ EF skills in a group context are not significantly affected by increases in small group size when there are between 3-7 children in the group. This runs counter to evidence from previous studies suggesting a decrease in attention and an increase in off-task behavior as a result of larger class sizes (Finn et al., 2003; Blatchford et al., 2003; Schanzenbach, 2014). Our findings indicate that decreasing the size of peer groups in small group activities may not affect kindergarteners' EF skills in the same way decreasing the total class size can.

Our next goal was to determine whether or not there was an interaction between group size and gender for predicting performance on the group-based EF measures. We anticipated that boys would be more affected by the size of the group than girls. This would result in boys performing worse than girls across all of the group-based EF measures, particularly when the size of the group was large. However, our results did not support this hypothesis since there was no significant interaction between group size and gender for predicting performance on any of the group-based EF measures. Therefore, the main effect of group size on the group-based measures was not significant regardless of gender.

The Role of Gender in Group-Based EF

The second small-group characteristic focused on in this study was gender ratio measured as the proportion of boys in the group. Our results show no significant main effect of proportion of boys in the group on any of the group-based EF measures, and no significant correlations between proportion of boys and performance on any of these measures. From prior research demonstrating robust gender differences in lab-based EF measures, teacher reports, and classroom observations (Matthews et al., 2009; Duckworth & Seligman, 2006), we expected that an increase in the proportion of boys in the group would result in kindergarteners performing worse on the group-based EF measures. Our results did not support this hypothesis. This provides evidence against a potential group dynamic of a higher proportion of kindergarten boys in a small group producing more distractions and decreasing overall attention. Therefore, this may not explain why boys underperform on ecologically sensitive measures such as CBRS and the *Gift Wrap* task while still producing strong results on lab-based achievement measures (Duckworth & Seligman, 2006; Matthews et al., 2009; McCabe & Brooks-Gunn, 2007).

In addition to running regressions to test for a main effect of proportion of boys in the group on the group-based EF measures, we checked for an interaction between proportion of boys and gender for predicting performance on each of the group-based EF measures. We expected that boys would be more susceptible to decreased EF performance on the group-based tasks than girls as a result of a larger proportion of boys in the group. However, we did not find a significant interaction between proportion of boys and gender for predicting performance on any of the group-based EF measures.

Although determining the main effect of gender on the group-based EF measures was not a primary goal of this study, in attempting to identify potential interactions between the small-group characteristics and gender, we did find significant main effects of gender on two group-based EF measures. Girls significantly outperformed boys on the group-based measures for attention and working memory (the distractibility and working memory composites). However, girls did not significantly outperform boys on the corresponding lab-based measures (Pair Cancellation and Backward Digit Span). On the other hand, the reverse of this phenomenon was true for the measures of response inhibition. Girls did not significantly outperform boys on the group-based measure of response inhibition (stop-time composite) but did significantly outperform them on the lab-based measure (HTKS), replicating previous findings of gender differences in HTKS (Matthews et al., 2009). These results suggest that there may be gender differences in EF skills depending on the setting in which they are measured. It is possible that classroom or individualized testing environmental factors disrupt EF skills in one gender more than the other. Gender-specific environmental influences may explain some research suggesting that girls significantly outperform boys on classroom-based tasks like CBRS and the Gift Wrap task but not on lab-based achievement measures (Duckworth & Seligman, 2006; Matthews et al., 2009; McCabe & Brooks-Gunn, 2007). However, more research is required to determine whether or not this is the case.

Limitations

Our study had several limitations worth considering when interpreting our findings. An important limitation to consider is how groups were selected for the group-based EF measures. Due to many factors including time constraints fitting students in for tasks, students could not be randomly assigned to groups. Groups were selected based on the availability of each student at a particular time in order to minimize interference with school activities. Without random assignment of groups, we cannot be certain that all potential confounding variables were accounted for. Furthermore, this resulted in groups of mostly four or five students since this was the amount of students that was most often available at any one time. A limited number of groups of three and groups of seven caused us difficulties in attempting to determine a significant difference between the smallest and largest groups in the sample. Our sample of groups also could not be manipulated to incrementally increase the number of boys or girls in groups to have equal numbers of groups of each gender ratio. Thus, many of the groups were between 33% and 66% boys which did not allow for a very thorough study of groups with more towards 100% of one gender or the other. This lack of random selection of groups for both of our main research questions could have influenced our results.

Despite the limitations of our study, our findings have the potential to help inform future research into the selection of small groups for elementary school activities. Given the importance of EF for academic achievement, choosing small groups in a way that optimizes kindergarteners' EF skills may be a strong alternative to ability grouping. However, more research in this area is required to determine whether or not this is the case. In our study, we expected that the smaller the group of students participating in demanding EF tasks, the better students would perform. However, it appears from our findings that the size of a group between 3-7 students does not significantly affect the manifestation of EF skills in these tasks. Therefore, preliminary findings suggest that kindergarten teachers may not need to be concerned with the size of small groups and may have more flexibility in selecting groups based on other factors. For our second research question, we anticipated that groups with a higher proportion of boys would produce more distractions, thus decreasing students' performance on demanding EF tasks. Since our results do not support this hypothesis, preliminary findings suggest that kindergarten teachers may be able to form small groups without consideration of the gender ratio of each group. Although this process of eliminating the concerns of group size and gender ratio when selecting small groups provides flexibility to teachers, without a more effective approach towards small-group selection, this information is not particularly useful. Future research studying peer effects on EF is essential for uncovering an approach with the potential for significantly enhancing kindergarteners' EF skills during small-group learning activities.

Conclusion

The present study was intended to examine the effects of the small-group characteristics, group size and gender ratio, on new group-based EF measures. No significant effects of group size or gender ratio on these measures were found. These results also did not depend on the gender of the child, as shown by the lack of interactions between these variables and gender in predicting group-based EF measures. However, these analyses did reveal significant main effects of gender on group-based EF measures of attention and working memory, which are worth exploring further. Overall, our findings provide insight into small group selection in kindergarten and serve as a step towards further research into the effects of small-group composition on EF skills in early elementary school children.

References

Ahmed, S.F., (2019). *Measuring Executive Function During Early Childhood: The Utility of Direct Assessments, Teacher Ratings, and Group-Based Tasks* (Publication No. 0000-0003-3814-2955) [Doctoral dissertation, University of Michigan].

Ahmed, S. F., Tang, S., Waters, N. E., & Davis-Kean, P. (2018). Executive function and academic achievement: Longitudinal relations from early childhood to adolescence. *Journal of Educational Psychology*. doi: 10.1037/edu0000296

Blatchford, P., Edmonds, S. and Martin, C. (2003), Class size, pupil attentiveness and peer relations. *British Journal of Educational Psychology*, *73*: 15-36. doi:10.1348/000709903762869897

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647-663. doi:10.1111/j.1467-8624.2007.01019.x

Bronson, M. B., Tivnan, T., & Seppanen, P. S. (1995). Relations between teacher and classroom activity variables and the classroom behaviors of pre-kindergarten children in Chapter 1 funded programs. *Journal of Applied Developmental Psychology*, *16*, 253–282.

Chorzempa, B. F., & Graham, S. (2006). Primary-Grade Teachers' Use of Within-Class Ability Grouping in Reading. *Journal of Educational Psychology*, 98, 529-541.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959–964. https://doi.org/10.1126/science.1204529

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2014). Self-control in school-age children. *Educational Psychologist*, *49*(3), 199–217. https://doi.org/10.1080/00461520.2014.926225

Duckworth, A., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98,* 198–208. doi:10.1037/0022-0663.98.1.198

Finn, J., Pannozzo, G., & Achilles, C. (2003). The "Why's" of Class Size: Student Behavior in Small Classes. *Review of Educational Research*, *73*(3), 321-368.

Kutnick P., Blatchford P. (2014) Groups and Classrooms. In: Effective Group Work in Primary School Classrooms. *Professional Learning and Development in Schools and Higher Education*, *vol 8*. Springer, Dordrecht

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-Class Grouping: A Meta-Analysis. *Review of Educational Research*, 66(4), 423–458. https://doi.org/10.3102/00346543066004423

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

McCabe, L. A., & Brooks-Gunn, J. (2007). With a little help from my friends? Self- regulation in groups of young children. *Infant Mental Health Journal, 28,* 584–605. doi.org/10.1002/imhj.20155

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early*

Childhood Research Quarterly, *21*(4), 471–490. https://doi.org/10.1016/j.ecresq.2006.09.003

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007b). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947–959. https://doi.org/10.1037/0012-1649.43.4.947

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

McClelland, M. M., Cameron, C. E., Wanless, S. B., Murray, A., Saracho, O. N., & Spodek, B. (2007a). Executive function, behavioral self-regulation, and social-emotional competence. *Contemporary Perspectives on Social Learning in Early Childhood Education*, *1*, 113-137.

McCoy, D. C. (2019). Measuring young children's executive function and self-regulation in classrooms and other real-world settings. *Clinical Child and Family Psychology Review*. https://doi.org/10.1007/s10567-019-00285-1

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

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety.

Proceedings of the National Academy of Sciences, USA, *108,* 2693–2698. doi:10.1073/pnas.1010076108

Moffett, L., & Morrison, F. J. (2019). Off-task behavior in kindergarten: Relations to executive function and academic achievement. *Journal of Educational Psychology*. https://doi-org.proxy.lib.umich.edu/10.1037/edu0000397

Morrison, F. J., Ponitz, C. C., & McClelland, M. M. (2010). Self-regulation and academic achievement in the transition to school. In S. D. Calkins, M. A. Bell, S. D. Calkins, M. A. Bell (Eds.), *Child development at the intersection of emotion and cognition* (pp. 203-224). Washington, DC, US: American Psychological Association. doi:10.1037/12059-011

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children*, *5*, 2, 113- 127.

Nesbitt, K. T., Baker-Ward, L., & Willoughby, M. T. (2013). Executive function mediates socio-economic and racial differences in early academic achievement. *Early Childhood Research Quarterly*, *28*(4), 774-783. doi:10.1016/j.ecresq.2013.07.005

Obradović, J., Sulik, M. J., Finch, J. E., & Tirado-Strayer, N. (2017). Assessing students' executive functions in the classroom: Validating a scalable group-based procedure. *Journal of Applied Developmental Psychology*, *55*, 4-13. doi.org/10.1016/j.appdev.2017.03.003

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Baltimore, MD: Paul H Brookes Publishing.

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

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, *45*(4), 958–972. https://doi.org/10.1037/a0015861

Schanzenbach, D. W. (2014). *Does class size matter?* Boulder, CO: National Education Policy Center.

Sørensen, A. B., & Hallinan, M. T. (1986). Effects of Ability Grouping on Growth in Academic Achievement. *American Educational Research Journal*, 23(4), 519–542. https://doi.org/10.3102/00028312023004519

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43*, 352–360. doi:10.3102/0013189X14553660

Wechsler, D. (1991). Manual for the Wechsler intelligence scale for children-(WISC-III). *San Antonio, TX: Psychological Corporation*.

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. https://doi.org/10.1016/j.ecresq.2012.12.002

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*(3), 436–452. doi:10.1016/j.jecp.2010.08.008

Wilkinson, I. A. G., & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research*, 37, 425–447.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock– Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Examiner's manual. *Woodcock-Johnson III Tests of Achievement*.

Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education (NCER 2017–2000)*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Research

Table 1

Means, Standard Deviations, Sample Size, and Range for Variables of Interest

Variable	N	M	SD	Min	Max
Gender (0 = Male)	252	0.47	0.50	0	1
Age at Group Games Testing	247	5.79	0.37	4.99	6.98
Group Size	251	4.81	0.81	3	7
Proportion of Boys in Group	251	0.54	0.17	0	1
Classroom-Based EF Tasks					
Stop-Time Composite	251	12388.5	3986.5	4526	33650
Distractibility Composite	251	36.8	8.2	17	79
Working Memory Composite	251	10.2	8.8	-37	24
Laboratory EF Tasks					
Head-Toes-Knees-Shoulders	249	31.0	16.8	0	58
Backward Digit Span	248	2.1	1.6	0	8
Pair-Cancellation	246	18.4	9.6	0	43

Table 2

Correlations

	1	2	3	4	5	6	7	8	9	10
1. Gender (0 = Male)	-									
2. Age at Testing	-.06	-								
3. Backward Digit Span	-.03	.12	-							
4. HTKS	.21***	.16**	.41***	-						
5. Pair Cancellation	.11	.25***	.34***	.31***	-					
6. Stop-Time Composite	.11	.16*	.26***	.28***	.14*	-				
7. Dist. Composite	.24***	-.04	.23***	.24***	.17**	.39***	-			
8. WM Composite	.15*	.20**	.31***	.18**	.25***	.26***	.21**	-		
9. Group Size	-.09	.22**	.11	-.02	.07	-.03	-.04	.15*	-	
10. Proportion of Boys	-.28***	.05	.15*	.02	.01	-.01	-.10	-.02	.13*	-

Note. HTKS = Head-Toes-Knees-Shoulders. Dist. Composite = Distractibility Composite.

WM Composite = Working Memory Composite.

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

Table 3

Stepwise Regressions of Group Size Predicting Stop-Time Composite (N = 234)

	Stop-Time Composite							
	Model 1				Model 2			
	$R^2 = .128$				$R^2 = .130$			
	B	SE	β	*p*	B	SE	β	*p*
School Attending	434.47	155.58	.179**	.006	440.23	155.98	.182**	.005
Age at Testing	1075.79	705.08	.100	.128	1055.70	706.48	.098	.136
HTKS	62.56	15.04	.264***	.000	62.24	15.09	.267***	.000
Group Size	-56.08	320.59	-.011	.861	-37.28	322.12	-.008	.908
Gender	427.37	500.85	.053	.394	417.15	501.64	.052	.406
Group Size x Gender					-415.28	606.37	-.042	.494

Note. HTKS = Head-Toes-Knees-Shoulders.

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

Table 4

Stepwise Regressions of Group Size Predicting Distractibility Composite (N = 234)

	Distractibility Composite							
	Model 1				Model 2			
	R^2 = .080				R^2 = .082			
	B	SE	β	*p*	B	SE	β	*p*
School Attending	0.18	0.33	.036	.597	0.17	0.33	.034	.611
Age at Testing	-1.61	1.51	-.074	.287	-1.61	1.51	-.074	.288
Pair Cancellation	0.15	0.06	.174**	.009	0.15	0.06	.175**	.009
Group Size	-0.19	0.68	-.019	.782	-0.21	0.68	-.020	.763
Gender	3.30	1.04	.201**	.002	3.29	1.05	.201**	.002
Group Size x Gender					0.68	1.28	.033	.598

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

Table 5

Stepwise Regressions of Group Size Predicting Working Memory Composite (N = 234)

	Working Memory Composite							
	Model 1				Model 2			
	R^2 = .156				R^2 = .157			
	B	SE	β	*p*	B	SE	β	*p*
School Attending	0.07	0.35	.013	.844	0.08	0.36	.015	.816
Age at Testing	3.93	1.56	.160*	.012	3.90	1.56	.159*	.013
Backward Digit Span	1.62	0.35	.286***	.000	1.63	0.35	.287***	.000
Group Size	1.03	0.73	.090	.158	1.06	0.73	.093	.148
Gender	3.24	1.11	.175**	.004	3.23	1.11	.175**	.004
Group Size x Gender					-0.78	1.37	-.034	.568

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

Table 6

Stepwise Regressions of Proportion of Boys Predicting Stop-Time Composite (N = 234)

	Stop-Time Composite							
	Model 1				Model 2			
	R^2 = .128				R^2 =.130			
	B	SE	β	*p*	B	SE	β	*p*
School Attending	442.98	150.54	.183**	.004	447.95	150.87	.185**	.003
Age at Testing	1039.16	678.81	.096	.127	1008.62	680.93	.094	.140
HTKS	62.61	15.05	.264***	.000	61.74	15.12	.260***	.000
Proportion of Boys	185.83	1537.66	.008	.904	373.81	1562.50	.015	.811
Gender	448.29	518.22	.056	.388	457.26	518.93	.057	.379
Proportion of Boys x Gender					-2173.74	3100.58	-.043	.484

Note. HTKS = Head-Toes-Knees-Shoulders.

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

Table 7

Stepwise Regressions of Proportion of Boys Predicting Distractibility Composite (N = 234)

	Distractibility Composite							
	Model 1				Model 2			
	R^2 = .082				R^2 = .082			
	B	SE	β	*p*	B	SE	β	*p*
School Attending	0.18	0.32	.037	.573	0.17	0.32	.035	.596
Age at Testing	-1.66	1.46	-.076	.255	-1.63	1.46	-.075	.267
Pair Cancellation	0.15	0.06	.175**	.009	0.15	0.06	.174**	.009
Proportion of Boys	-2.04	3.29	-.040	.536	-2.18	3.32	-.043	.512
Gender	3.15	1.07	.192**	.004	3.15	1.08	.192**	.004
Proportion of Boys x Gender					2.39	6.60	.023	.718

* p < .05, **p < .01, ***p < .001

Table 8

Stepwise Regressions of Proportion of Boys Predicting Working Memory Composite (N = 234)

	Working Memory Composite							
	Model 1				Model 2			
	R^2 = .150				R^2 = .155			
	B	SE	β	*p*	B	SE	β	*p*
School Attending	-0.06	0.35	-.011	.856	-0.08	.345	-.015	.807
Age at Testing	4.53	1.51	.184**	.003	4.68	1.52	.190**	.002
Backward Digit Span	1.66	0.35	.292***	.000	1.68	.352	.296***	.000
Proportion of Boys	-1.65	3.51	-.030	.639	-2.38	3.55	-.043	.504
Gender	2.96	1.15	.161*	.010	2.95	1.15	.160*	.011
Proportion of Boys x Gender					8.62	7.02	.075	.221

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