Investigation of Strategies to Reduce the Impact of the Relative Age Effect in Kindergarten

NCT05142826

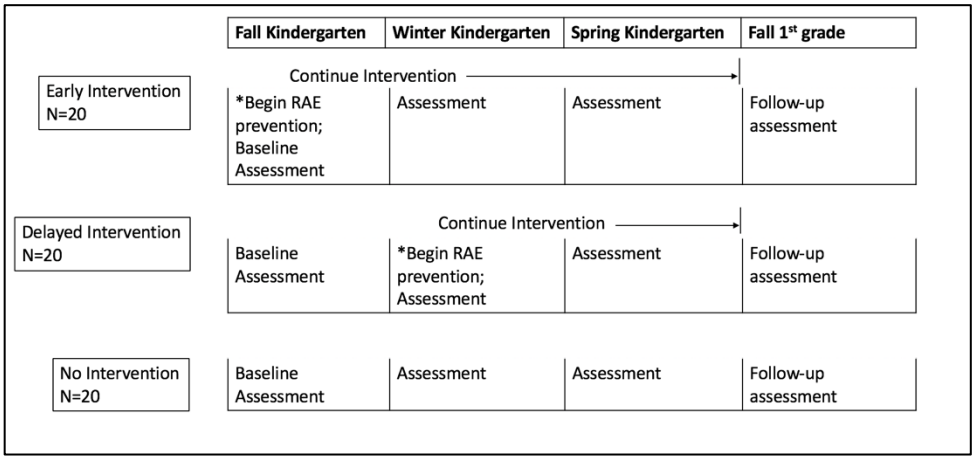
Document Date: November 10, 2023

Initial Pilot Intervention Study

Following development of all intervention materials, in the spring of the first year we will conduct a multiple baseline design across at least three classrooms to investigate the preliminary effectiveness of our approach to promoting positive student outcomes. In the spring of the first year, once initial intervention materials are developed, we will begin collection of concurrent baseline data for a child with ADHD and young for the grade from at least three participating classrooms. Then, one classroom will be assigned to begin intervention based on an analysis of baseline stability, level, variability, and trend and subsequent classrooms will be introduced to intervention using the same criteria at later timepoints. Thus, the project will be a concurrent multiple baseline design across classrooms, with intervention implemented in systematic steps across classrooms at varied timepoints. This will allow for an evaluation of the initial scientific premise of the developed intervention, and provide a context for the investigators to assess implementation of treatment, timely collection of measures, feasibility of procedures, and collect information on teacher satisfaction and suggestions for modification. The scientific rigor of this pilot study will be supported through an adherence to What Works Clearinghouse Standards for single-case design research (What Works Clearinghouse, 2019).

Specifically, teachers will be asked to complete a direct behavior rating (DBR) each day during a specified time period where the child exhibits challenging behavior. DBRs are brief ratings of student behavior (Briesch et al., 2016; Chafouleas, Riley-Tillman, & Sugai, 2007; Fabiano et al., 2017). DBRs will serve as the ongoing, progress monitoring measure in this pilot study. For example, following group instruction in reading, the teacher might circle a number on a scale from 0-10 to indicate the degree to which the student was actively engaged. The study will use the most studied target behaviors - Academic Engagement (AE), Disruptive Behavior (DB), and Respectful Behavior (RS) at the end of each targeted intervention session for each student. Inter-rater agreement will be collected on the DBRs by having an observer in the classroom for at least 20% of days for the same time period and completing identical DBR-SIS ratings as well as the SBTR observation. Teachers will complete all measures at initial baseline data collection, and at the end of the treatment. Child-based neurocognitive measures will be completed at the initial baseline data collection period and at the end of treatment. Thus, we will gain information on the feasibility of data collection to inform our larger, between group pilot study during this initial pilot as well.

Pilot Study of RAE Prevention Timing and Effectiveness



To explore the relationship between the hypothesized mechanisms of treatment and key outcomes, a small randomized trial will be conducted to investigate the feasibility and promise of the proposed treatment for modifying the hypothesized treatment targets. It is currently unclear *when* the RAE prevention should be optimally implemented. On the one hand, it may be beneficial to work with teachers immediately upon school entry, prior to the development of frequent relative comparisons that disadvantage the relatively young child in the class with ADHD. On the other hand, one might speculate that the intervention may be better positioned to promote teacher change after they have had frequent experiences with the child, and the relative comparisons have occurred, providing an active target for modification through intervention. Teacher acceptance and intervention feasibility may also vary based on implementation timing. Therefore, a pilot trial will be conducted that can compare the feasibility, effectiveness, and acceptability of these two competing intervention initiation timepoints, compared to a no intervention condition.

The sample of kindergarteners will be drawn from the school districts in Erie County. During the summer preceding each academic year, the UB researchers will partner with the school districts to attend kindergarten screening visits. All children who will enroll in kindergarten for the upcoming fall must come to the district office or designated school to enroll in the kindergarten year. We will partner with district administrators to provide time for study staff to introduce the study and request parental consent for participation. We will include all children born between 8/1 and 11/30 of the current school year. Specifically, children born in the

months of August, September, October, or November will be designated as relatively young for kindergarten. The informed consent process will include a discussion of the risks/benefits of the planned study approach, and inform parents of their alternative choice (enroll in kindergarten as usual). Our enrollment plan is feasible given the overall level of kindergarten enrollment each year. We expect to enroll 30 participants per year.

Inclusion criteria will include a child who is: (1) enrolling in kindergarten for the first time; (2) has a date of birth is between 8/1 and 11/30 of the current academic year; (3) does not have an Individualized Education program for behavioral concerns; and (4) has elevated symptom ratings of ADHD. Excluded will be any child outside of the targeted birthdate range, currently in restrictive special education placements for behavior (as this is a targeted study outcome), or taking psychoactive medication for mood, behavior, or inattention.

During kindergarten screening we will include children who have at least six symptoms endorsed at clinically significant levels via parent and/or preschool teacher ratings on the DBD rating scale across inattentive and/or hyperactive/impulsive presentations (Pelham, Gnagy, Greenslade, & Milich, 1992). Because we are exploring whether symptom ratings worsen during the school year, based on group assignment, we do not require a research-based, DSM diagnosis of ADHD, but expect our sample to illustrate symptom levels that place the child at risk for kindergarten classroom impairment due to ADHD. This approach also includes students who will eventually meet criteria for combined presentation, but exhibit primarily hyperactive/impulsive symptoms due to the lack of attentional demands in preschool/child care settings (Lahey, et al., 2005). Random assignment will be stratified by sex to distribute proportionally across groups. The project will purposely only include children born in the last quarter of the school year, as this is the key demographic of interest in the present proposal.

Children whose parents provide permission will be enrolled in (1) kindergarten as usual; (2) the RAE prevention intervention implemented immediately (i.e., in October following the assessment); or (3) the RAE prevention intervention implemented at mid-year (i.e., in February). Given variability in birthdays, the large number of kindergarten classrooms currently in the are targeted, and the modest study demands for parents, we expect to meet all recruitment targets outlined in this application. Children assigned to the RAE prevention groups will have a school psychologist consultant assigned to the student for the duration of the school year. The consultant will work with the child's primary teacher to provide the preliminary training on the RAE and positive classroom management strategies, and to establish a DRC over the course of two consultation visits.

For the positive behavioral support, target behaviors will be identified, operationally defined, and integrated into a DRC. The teacher will be asked to implement the intervention during the week between the first and second visit. At the second consultant visit, target behaviors will be refined, and using the data collected by the teacher, criteria for each target behavior will be modified (e.g., a child who averaged 10 verbally intrusive behaviors per class would have a target behavior changed to "Has 8 or fewer verbally intrusive behaviors"). DRCs will be implemented across settings and classes; however, a single teacher will be identified for assessment completion, observations will be conducted in that teacher's class. Parents will be invited to the school meetings, and the school psychologist will work with the parent to establish a menu of home-based privileges (e.g., screen time, bed time, extra parent-child activity time) that will be rewarded based on DRC performance each day. At the end of each day, the teacher will send the DRC home with the child so that the parent is provided feedback on a daily basis regarding the child's behavior at school. In addition to the home-based contingency management based on school feedback, which makes the child accountable at home for school-based behavior, the DRC serves as a mechanism of daily communication between the parent and teacher. Given the variability in behavior expected with a child with ADHD, such daily communication is often crucial for preventing deterioration in functioning and presents the parent and teacher with multiple opportunities to problem-solve.

After the initial meetings with the child's teacher, consultants will meet monthly with the teacher (and parent if available) via phone or video conference to provide feedback on the child's behavior during the month using a graphical representation of the child's DRC performance. This information will be used for data-driven decision making, a procedure shown to be better than teacher impressions/ judgments when intervention planning/monitoring (Fuchs et al., 2000). DRC goals will be adjusted as needed. It is expected that these monthly meetings will serve the purpose of ensuring goals are consistently addressed on the DRC throughout the year and that the goals set are developmentally appropriate for the child.

Kindergarten as Usual (KAU) group. Consultants will conduct an initial meeting with each teacher of children in the KAU group to describe study procedures and establish regular classroom observation times. This condition is conceptualized as a kindergarten as usual control and serves as the counter-factual condition.

Consultants will be M.A. level school psychology specialists or special education teachers appointed to the project. We have chosen to use individuals with these credentials to increase the external validity of the

intervention as these individuals are likely to be called upon to consult with teachers regarding the behavior of children with ADHD as well as be involved in the assessment of children with ADHD in many school districts.

Outcome Measures.

Outcome measures will be collected at the beginning of the kindergarten year in October, the middle of the kindergarten year in January, at the end of the kindergarten year in May, and at the beginning of the 1st grade year in October. These four assessment points will provide data on the proximal outcomes of the study manipulation and distal outcomes of the manipulation, and they will double as an assessment of the feasibility of the team's measurement plan and ability to retain participants in a longitudinal design (for both treatment and control groups).

ADHD-related Outcomes

Observations of Classroom behavior. Independent observations will be conducted using the Student-Behavior Teacher-Response observation system (SBTR; Pelham, Greiner, & Gnagy, 2008; Vujnovic, Fabiano et al., 2014). The SBTR is an observation code that collects information on: (1) the frequency of student rule violations; (2) whether a teacher observed the misbehavior; (3) if observed, the teacher consequence employed and whether it was an appropriate consequence. The SBTR system also records the number of praise statements and commands issued. The SBTR is a well-defined and validated observation system for use with children with ADHD in classroom settings (Pelham, Fabiano, & Massetti, 2005). The SBTR has been successfully employed as a primary outcome in similar projects (Fabiano, Vujnovic et al., 2010; Fabiano et al., 2013; Pelham, Fabiano et al., 2016). Each observation will last for thirty minutes. Time of day and activity will be controlled. Observers will be blind to the child's treatment condition. To ensure the inter-rater reliability of this measure, 20% of observations will be completed simultaneously with a second observer.

Disruptive Behavior Disorders (DBD) Rating Scale. ADHD Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 2000) symptoms will be measured using the DBD rating scale (Pelham et al., 1992), which will administered to the child's teacher and parent. The DBD is a 45-item measure that asks parents to rate the DSM symptoms of ADHD, ODD, and CD on a 0-3 point frequency rating scale (i.e., *Not at all, Just a little, Pretty Much, or Very* Much). The 0-3 ratings will be summed across 18 ADHD symptoms and then an average score for ratings across symptoms will be computed for each individual.

Impairment Rating Scale (IRS). The IRS (Evans, Allen, Moore, & Strauss, 2005; Fabiano et al., 2006) is a rating scale that asks teachers to rate the severity of the child's problems and need for treatment and/or special services in important functional domains (i.e., relationship with peers, relationship with the teacher(s), academic progress, classroom functioning, self-esteem, overall need). There are six items on the scale, and scores on the measure range from 0 (Not a problem/Definitely does not need treatment or special services) to 6 (Extreme problem/Definitely needs treatment or special services). The average score for each of the individual domains rated were used in the analysis; scores could range from 0.0-6.0.

Academic Functioning

Academic Performance Rating Scale (APRS). The APRS is a 22-item measure that asks teachers to rate different aspects of a child's academic performance in the classroom (DuPaul, Rapport, & Perriello, 1991). The APRS has three factors (DuPaul et al., 1991), two of which will be used in the present investigation: an Academic Productivity factor (7 items) that relates to work completion and accuracy, independent work, and following directions, and an Academic Success (12 items) factor related to the child's academic ability.

Pilot measures

Although dichotomous outcomes are unlikely to have enough variability in the present pilot studies, we will work to establish procedures to reliably collect this information within the current participants, in anticipation of collecting these data in a larger, future, clinical trial.

Referral for special education. All referrals from kindergarten and first grade for special education placement will be documented by the participating schools.

Stimulant medication use. Parents will be asked about the use of stimulant medication at any time during the school year.

Grade retention. At the end of each year, each participating child will be coded as being promoted or retained. This outcome will be assessed at the end of the kindergarten and 1st grade.

Assessment of Hypothesized Mechanisms of Treatment Effect

Positive Student-Teacher Interactions. The ratio of positive to negative statements issued by the teacher on the SBTR observational measure will be used as an indicator of student-teacher interactions, with

classrooms that have a positive environment having a larger number (i.e., more positives than negatives) relative to those that have a negative environment (i.e., fewer positives than negatives). Prior studies have demonstrated this information can be reliably collected (Vujnovic, et al., 2014). Further, these rates of positive to negative statements can be reliably modified based on background behavioral interventions such as the DRC (Fabiano et al., 2007). This measure will be collected across the three groups.

Executive Functioning. EF is thought to be a key construct in young children (e.g., Diamond & Lee, 2011). Despite the promise of studying EF in early development, researchers seeking to assess EF in early childhood face a number of conceptual and methodological problems (see, e.g., Best & Miller, 2010; Garon et al., 2008). Although numerous tasks have been employed to measure EF in early childhood (e.g., Carlson, 2005), many exhibit either ceiling or floor effects in the age range of the proposed study, use pass/fail rules that make them relatively insensitive, have too few items to reliably assess the construct of interest, and/or are conceptually "dirty", making it difficult to know what process is being assessed (see, e.g., Best & Miller, 2010; Carlson, 2005; Garon et al., 2008; Willoughby et al., 2012). The proposed assessments are based in theory and empirical work that suggests that the great diversity of putative EFs likely build on three core functions: working memory, inhibitory control, and cognitive flexibility or set shifting (e.g., Garon et al., 2008; Miyake et al., 2000). *Working memory* is the capacity to maintain temporary, active representations of information and to manipulate that information, a function critical for planning, problem-solving, and decision-making (e.g., Baddeley 2003). *Inhibitory control* is itself multidimensional (e.g., Nigg, 2000), but here we focus on response inhibition, the ability to stop a typical or pre-potent response; in practical terms, this is the ability critical for refraining from 'blurting out' an answer to a teacher's question. *Set shifting* involves flexibility of attention and responses when the rules/contingencies change. Set shifting likely involves both working memory and inhibition, but differs in its focus on adapting to changing circumstances (e.g., Best & Miller, 2010).

We seek to assess each of the three core domains in a relatively pure manner with developmentally sensitive and appropriate computerized tasks. Importantly, the proposed tasks are sensitive to cross-sectional age differences in the developmental period studied here; the proposed within-subjects approach will provide even greater power to detect developmental changes (e.g., Best & Miller, 2010; Carlson, 2005; Tillman et al., 2008). Ideally, we would assess each construct with multiple measures (e.g., Campbell & Fiske, 1959), but the additional time required to do so could seriously threaten construct validity within this young sample. Although we assess each domain separately, we will also consider constructing a single task-EF composite, given evidence that EF may be unidimensional in this age range and that such composites may exhibit better criterion validity than each separate index (e.g., Willoughby et al., 2012). We have had great success in collecting complete data from young children, including those with ADHD and other behavior problems, in more lengthy protocols for several consecutive days (e.g., Ashare, Hawk, et al., 2010; Bubnik et al., 2015; Hawk et al., 2018; Shiels, Hawk, et al., 2009; Spencer, Hawk, et al., 2009; Strand, Hawk et al., 2012). Task order is counterbalanced across participants.

All EF measures will be assessed at all four time points. This frame allows the maximal reasonable timeframe for assessment within the Kindergarten year (avoiding transition months of September and June). The Working Memory and Inhibitory Control tasks described below have mediated stimulant medication effects in an analogue elementary school classroom (Hawk et al., 2018; Preliminary Study 3), making them viable candidates as mechanisms for change in our RAE-focused intervention.

Working Memory. The spatial span task is a computerized adaptation (see Shiels, Hawk, et al., 2008) of prior working memory tasks, the Corsi blocks, and the WISC-IV spatial span measure (Kaplan et al, 2004; Luciana, 2003; Milner, 1971). An array of ten squares is presented on the computer screen. On each trial, a ☺ appears in a sequence of 2-8 squares (1/sec). For forward span, participants use a mouse to click the squares in the same order. For backward span, which requires manipulation of information in WM, participants click the squares in the reverse order. In each version, the task begins with a practice followed by 2-location sequences. The task ends when both trials within a difficulty level are incorrect. The number of trials correct is the primary dependent measure, with trials forward representing WM storage and trials backwards providing an index of WM manipulation (see Shiels, Hawk, et al., 2008).

Inhibitory Control. Response inhibition will be assessed using the stop *task*. The stop-signal paradigm (Logan, 1994; Logan et al., 1984) offers the most pure measure of response inhibition (e.g., Nigg, 2001). Participants button press to indicate whether the "go" signal (← or →) is pointing left or right, but are asked to not respond on trials when the stop signal (a brief tone or a second visual stimulus) is presented (25% of trials after the go stimulus). The stop signal delay adjusts dynamically based on whether the previous stop trial led to successful inhibition (next stop delay is longer, making it harder to inhibit) or a failure to inhibit (next stop delay

is shorter, making it easier to inhibit; Logan et al., 1997). The primary dependent variable, stop signal reaction time or SSRT, is an estimate of the speed of inhibitory processing.

Set shifting. To assess sensitivity and flexibility to changing task requirements and rules, we will employ the dimensional change card sort (DCCS) of Frye et al. (1995). To avoid ceiling effects on performance, we will follow the standard DCCS, followed by the advanced Borders DCCS (Carlson, 2005; Hongwanishkul et al., 2005) using the published protocol of Zelazo (2006). After brief initial practice, the child sorts stimuli that vary by shape (rabbit, boat) and color (red, blue). First, stimuli are sorted based on one dimension. Next, the rule switches to the other dimension. Most children in the present sample are expected to perform quite well on the first two stages. The advanced Borders DCCS is the last phase. In the borders version, children sort based on one dimension (color) if the stimulus has a black border around it, but based on the other dimension is there is no border. The Borders version will be scored continuously with the number sorted correctly serving as the primary dependent variable. Total testing time for the DCCS is 10 minutes.

Measures of Integrity and Fidelity

To record the integrity and fidelity (Cordray & Pion, 2006; Gresham, 1989; Lane, Beebe-Frankenberger, Lambros, & Pierson, 2001; Waltz, Addis, Koerner, & Jacobson, 1993) with which the intervention is implemented, multiple procedures will be used.

Because date of birth is the primary independent variable in this study, we will verify each child's date of birth using the official Certificate of Live Birth required for kindergarten registration. This review will occur during the explanation of the informed consent document that is conducted at kindergarten screenings.

Data Analysis Plan

Specific aim one will be addressed through developmental activities to create and refine the intervention (see description of Deliverables above). This will include the creation of streaming videos, consumable materials, technical innovations to realize program components, and the manualized procedures. For the pilot, multiple baseline study, data will be analyzed by visual inspection to see whether slope and level of the measures change upon the initiation of the intervention and remain stable for the baseline phases (WWC, 2017). A key analysis will be an investigation of whether any changes in slope and level were dependent on phase change, and whether any such changes were replicated across cases. Across phases, the study investigators will consider the immediacy of effect, overlap, and consistency of data in similar phases (What Works Clearinghouse, 2019). We will collect teacher ratings of acceptability and observations of adherence.

Data analyses will also address specific aim two, "Investigate whether key academic and social outcomes are significantly improved in an intervention to prevent the RAE (immediately and after a delay) vs. KAU." For each of the outcome measures, data will be analyzed using analyses within SPSS. The analysis will be a 3 (Group: RAE Prevention Immediate, RAE Prevention Delay, KAU) by 4 (Time: Pre-treatment, Mid-year, End-of-year, First Grade Follow-up) Mixed Models analysis with group as the between group factor and time as the within subject factor. Also, using procedures outlined in Cohen (1992), effect sizes will be calculated to display the magnitude of treatment effect for the RAE programs relative to KAU (for dichotomous outcomes we will calculate odds ratios or relative risk). Effect sizes will be calculated for each outcome measure. Of particular interest are also variables related to establishing the mechanisms responsible for intervention effects and analysis of acceptability and implementation fidelity.

Most of the process measures collected in this study are descriptive and will help establish the feasibility of procedures for a larger R01 application. Thus, all descriptive information on adherence to study procedures and compliance with study assessments will be collected to evaluate the feasibility of this approach within a larger trial. Further, we will evaluate whether there are any negative side effects of the RAE prevention intervention, by inspecting whether teachers differentially rate children in the RAE groups as more symptomatic/impaired as the intervention is implemented (e.g., children may initially behave worse prior to getting better if the DRC is subjecting them to behavioral demands).

Statistical Power

We are purposely not using this pilot study as a mechanism to generate effect size estimates for an efficacy study given the clear caution in the field against this as a goal for pilot studies such as this one (see Kraemer et al., 2006; Leon, Davis, & Kramer, 2011). Rather, the pilot study will be prioritized as a way to determine the promise of the procedures to be implemented reliably and with integrity in authentic school settings. Further, we will explore whether the hypothesized mechanisms are related to study outcomes as described above.