School Psychology Review,
2014, Volume 43, No. 3, pp. 239–259

Enhancing the Academic Development of Shy Children: A Test of the Efficacy of INSIGHTS

Erin E. O'Connor, Elise Cappella, Meghan P. McCormick, and Sandee G. McClowry
New York University

Abstract. This study investigated the efficacy of the INSIGHTS into Children's Temperament intervention in supporting the academic development of shy kindergarten and first-grade children. INSIGHTS is a temperament-based intervention with teacher, parent, and classroom programs. The participants included 345 children from 22 low-income, urban elementary schools who were randomly assigned to INSIGHTS or a supplemental after-school reading program. Growth-curve modeling showed that shy children in INSIGHTS evidenced more rapid growth in critical thinking and math than their shy peers in the attention-control condition during kindergarten and the transition to first grade. The effects of INSIGHTS were partly indirect through improved behavioral engagement. INSIGHTS enhances the academic development of early elementary school children with shy temperaments.

Ensuring that all children are ready to learn in early elementary school remains a concern of educators, researchers, and interventionists. Academic success in the early elementary school years has a clear link to subsequent social, economic, and health outcomes (Heckman, 2006). Because of its importance, a large body of research has focused on identifying factors related to early academic skill development. Although the magnitude of associations has been modest, shy temperament is consistently related to early academic skill difficulties, with negative trajectories worsening across middle childhood (see Evans, 2010, for a review). It may be particularly important to examine this temperamental risk for low-income, urban children, already at risk for low academic achievement (Elias & Haynes, 2008).

A number of terms such as *socially withdrawn*, *isolated*, and *anxious-solitary* are used to describe children with shy temperaments, but they have common characteristics despite the term being used. In this study, we describe temperamentally shy children as being fearful, anxious, wary, and reluctant to take part in interactions with others in situations that involve uncertainty, novelty, and actual or perceived judgment by others (Evans, 2010). In accord with Evans (2010), we use the term *shy* when making generalizations across studies that examine the effects of these behavioral tendencies on academic skill development.

The properties of many classroom environments and teacher expectations are often not aligned with shy children's styles of behaving. As early as kindergarten, teachers expect their students to participate in classroom

Correspondence regarding this article should be addressed to Erin E. O'Connor, Department of Teaching and Learning, New York University, 239 Greene St, New York, NY 10003; e-mail: eoc2@nyu.edu

activities and to interact socially with their peers. Instead, shy children are reluctant to volunteer or answer questions and often are not assertive in their interactions with peers and adults (Rudasill & Konold, 2008).

The reticence of temperamentally shy children results in less behavioral engagement in the classroom compared with their peers. *Behavioral engagement* refers to students' effort, persistence, concentration, and interest in classroom activities (Skinner, Kindermann, & Furrer, 2009). Shy children's lower levels of behavioral engagement in the classroom are, in turn, linked with lower-level academic skills (Hughes & Coplan, 2010). Although temperament is relatively stable, behavioral engagement is malleable and responsive to intervention (Hughes & Coplan, 2010). Given the long-term consequences of early academic skill difficulties (Alexander, Entwisle, & Kabbani, 2001), interventions that support the behavioral engagement of shy children are necessary.

Temperament interventionists aim to enhance goodness of fit (McClowry & Collins, 2012), which is the match between the environment's demands, opportunities, and expectations and an individual's temperament (Chess & Thomas, 1984), to support individuals' positive engagement with their surroundings. Improving the responsivity of the environment is particularly important in early education because academic outcomes are more strongly associated with temperament in the early grades than with cognitive aptitude (Entwisle, Alexander, & Olson, 2005). Moreover, entering kindergarten and first grade are particularly critical transition times for children, which subsequently affect their educational attainment (Bossaert, Doumen, Buyse, & Verschueren, 2011). Responsive teacher and parenting strategies can be implemented to assist children in meeting the new environmental demands and supporting their positive engagement with the classroom (Curby, Rimm-Kaufman, & Ponitz, 2009).

SHYNESS AND ACADEMIC SKILLS DURING TRANSITION TO SCHOOL

A growing body of literature indicates that shy children evidence lower levels of academic skills than their more outgoing peers in early and later childhood. Spere and Evans (2009) found that at 5 to 7 years of age, shy children scored lower on tests of prereading skills and reading mastery than their more outgoing peers. Similar associations between shyness and lower-level reading scores have been identified among children in later elementary school (e.g., Rapport, Denney, Chung, & Hustance, 2001; Slomkowski, Nelson, Dunn, & Plomin, 1992).

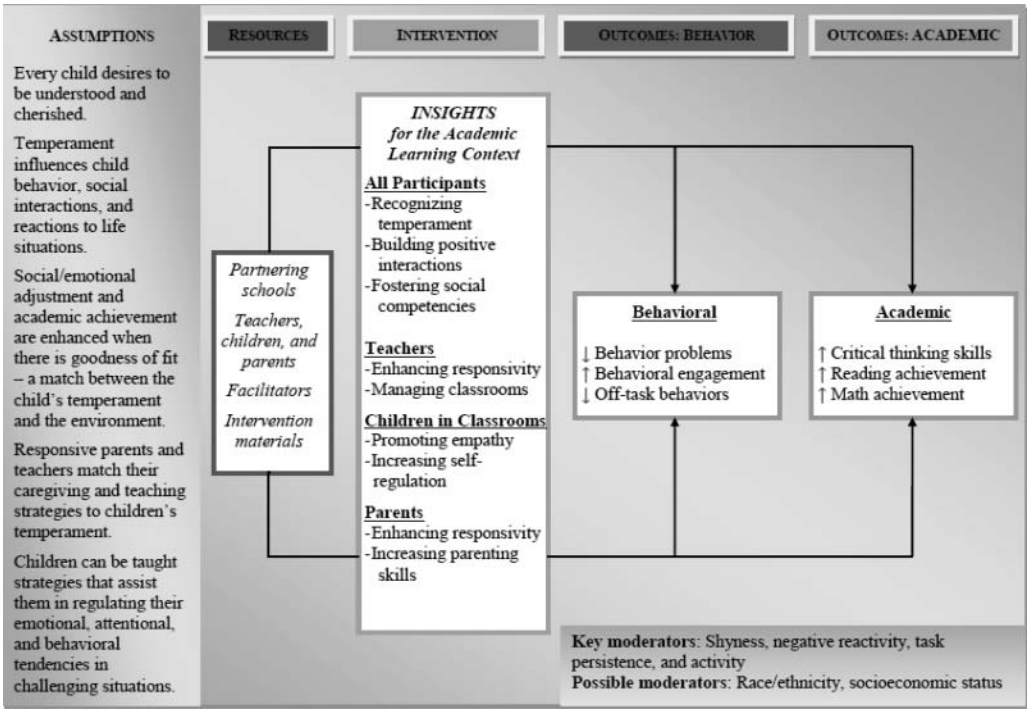
Shy children also evidence lower levels of math achievement. For instance, in a cross-sectional study of preschool children, Dobbs, Doctoroff, Fisher, and Arnold (2006) found that shy children scored significantly below their more outgoing peers on mathematics achievement tests. The associations between a shy temperament and lower levels of mathematics achievement persist into elementary school (e.g., Maziade, Côté, Boutin, Boudreault, & Thivierge, 1986). Although no empirical studies have been conducted to examine the effects of a shy temperament on critical-thinking skills, such associations would be expected to exist. Children who are shy become overwhelmed by situations that are emotionally charged (Denham & Brown, 2010). As a result, they may lack the personal resources to focus on activities that require critical-thinking skills.

SHYNESS AND BEHAVIORAL ENGAGEMENT: A POSSIBLE MEDIATING MECHANISM

Robust empirical literature links shyness to reticent behavior in the classroom on the first day of school (Coplan, 2000; Gersten, 1989) as well as several months into the school year (Coplan, Findlay, & Nelson, 2004; Rimm-Kaufman & Kagan, 2005). Shy children volunteer less often and speak less in class than their more outgoing peers (Evans, 1987; Gordon & Thomas, 1967). In addition, shy children are more likely to withdraw from social interactions with their classmates (e.g., Coplan et al., 2004; Ladd & Profilet, 1996).

Behavioral engagement is related to academic achievement because children learn

Figure 1. Logic Model for INSIGHTS Into Children's Temperament for Improving Academic Skills



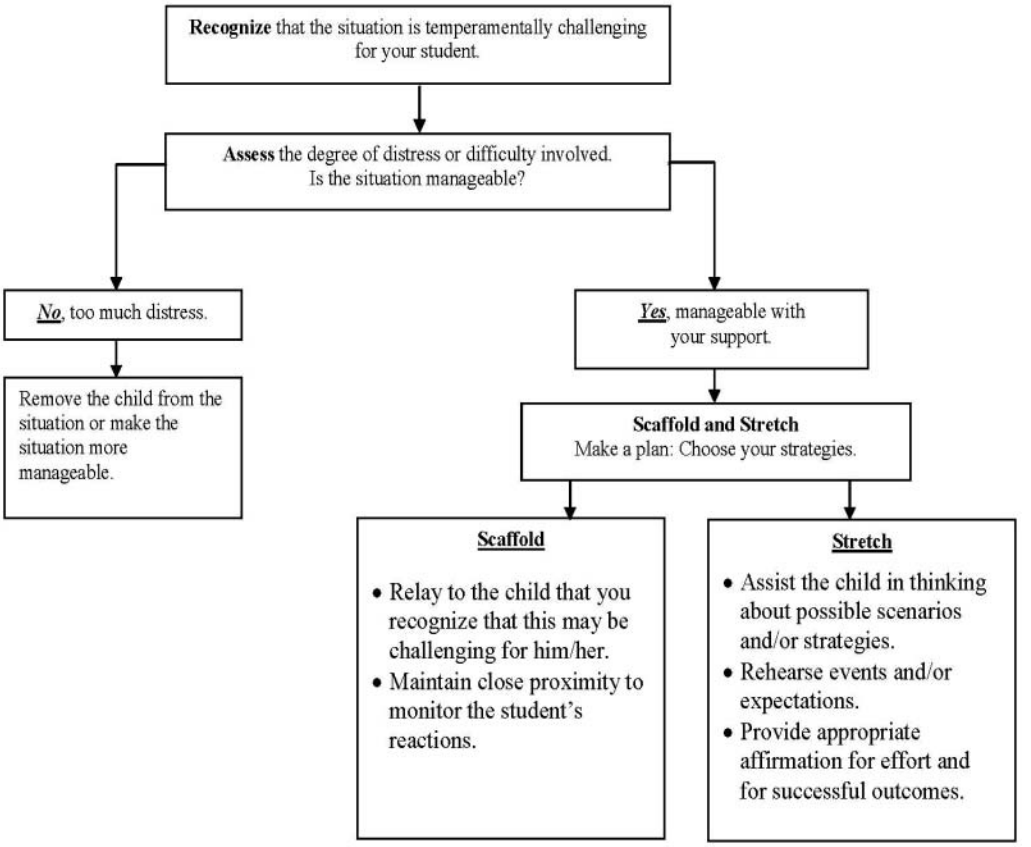
academic content and skills through their classroom involvement (Hughes & Coplan, 2010). Furthermore, perceived levels of behavioral engagement may influence teachers' perceptions of students' abilities, in turn leading teachers to rate students whom they perceive as less engaged as having lower-level academic skills than their more engaged peers (Maldonado-Carreno & Votruba-Drzal, 2011). Teachers tend to perceive shy students as unengaged because such children have a longer latency period before speaking and offer fewer verbal requests (Crozier & Perkins, 2002). As early as age 4 years, shy children are rated by their teachers as less intelligent and academically competent (Gordon & Thomas, 1967; McCroskey & Daly, 1976; Nelson, Rubin, & Fox, 2005; Richmond, Beatty, & Dyba, 1985).

Descriptive studies have found that behavioral engagement is a critical mechanism for the effects of shyness on academic skill development (Bruce, Davis, & Gunnar, 2002; Downer, Rimm-Kaufman, & Pianta, 2007; Fantuzzo et al., 2007; Hughes & Coplan, 2010). For example, results from a quasi-experimental study suggested that among children in Grades 4 to 6, behavioral engagement, in part, mediated associations between shyness and achievement (Hughes & Coplan, 2010). In other words, the poorer academic skills of shy children were partially accounted for by their lack of engagement in the classroom. Thus, enhancing the behavioral engagement of shy children may be a critical conduit for supporting their academic outcomes.

INSIGHTS INTO CHILDREN'S TEMPERAMENT INTERVENTION

The INSIGHTS into Children's Temperament intervention is an intervention designed to enhance the development of low-income primary-grade children at risk for academic and behavioral difficulties. As depicted in the logic model in Figure 1, INSIGHTS is a comprehensive temperament-

Figure 2. Scaffolding and Stretching



Note. Helpful steps to use when a student encounters a temperamentally challenging situation are shown.

based intervention for teachers, parents, and children that integrates theory, research, and clinical strategies intended to support the academic learning context. Using a temperament interventionist perspective, INSIGHTS aims to enhance goodness of fit. In the parent and teacher program sessions, caregivers learn to recognize a child's temperament based on four empirically derived temperament typologies (McClowry, 2002; McClowry et al., 2013): Hilary the Hard Worker, who is industrious; Gregory the Grumpy, who is high maintenance; Fredrico the Friendly, who is social and eager to try; and Coretta the Cautious, who is shy. Parents and teachers are then encouraged to reframe their perceptions more positively and to select strategies that match a child's particular temperament. For example, parents and teachers are taught to recognize that a child's temperament is shy and thus the child has a tendency to withdraw. Reframing disavows the premise that a shy child could be more social if only he or she tried. Instead, caregivers are encouraged to appreciate that shy children are often astute observers who are sensitive to their environment and cautious when encountering new situations or meeting people.

As illustrated in Figure 2, teachers and parents also learn to use a scaffold-and-stretch approach when children encounter temperamentally challenging situations, which are particularly relevant strategies for encouraging the behavioral engagement of shy children. For example, when assigning a classroom activity that is challenging to a shy child, such as

participating in a school assembly, teachers are encouraged to assess the student's distress level. If the child is likely to be unduly distressed, the teacher can make the activity more manageable. Although a young shy child might be overwhelmed by a major role in an assembly, a lesser role with more opportunities to rehearse could stretch the child from a reticent participant to one who is successfully engaged. To support the child, the teacher might first arrange a practice session with a friend and then with a small group of classmates before a whole-class rehearsal (McClowry, 2014). Acknowledging the child's efforts in each progressively challenging step is another part of this responsive strategy.

In INSIGHTS classroom sessions, activities focus on empathy and problem-solving skills. The students are introduced to puppets exemplifying the same four temperament typologies. The children are also encouraged to understand their respective strengths and challenges. For example, the Coretta the Cautious puppet thinks carefully before she acts but warms up when provided more time. The children also use the puppets to resolve videotaped dilemmas and those they encounter in their own lives. In a dilemma involving the Coretta the Cautious puppet, the children and the puppets encourage her to be assertive when she hesitates to ask her teacher for help.

In a previous prevention trial (McClowry, Snow, & Tamis-LeMonda, 2005), INSIGHTS was efficacious in reducing children's disruptive behaviors, especially among children who were at diagnostic levels of one or more disruptive behavior disorders, including attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. A second study (O'Connor et al., 2012) showed that INSIGHTS enhanced parenting efficacy, especially among parents whose children had high-maintenance temperaments that were low in task persistence and high in negative reactivity and activity.

The previous studies did not examine the intervention's impact on academic outcomes or differential program effectiveness for shy children, a gap this analysis from the third randomized controlled trial was designed to fill. INSIGHTS was expected to be beneficial for shy children whose heightened sensitivity often compromises their academic outcomes (Hughes & Coplan, 2010). By using temperament-responsive strategies, caregivers would improve the fit between shy children and their environments. The children were also anticipated to apply strategies they learned to enhance their own self-regulation in temperamentally challenging situations, thereby supporting their own behavioral engagement in the classroom.

PURPOSE

The purpose of the current analysis was to investigate the efficacy of INSIGHTS in supporting the academic development (critical-thinking, math, and language skills) of shy children in urban, low-income schools during kindergarten and across the transition into elementary school. In addition, we examined whether behavioral engagement mediated associations between INSIGHTS and academic development for shy children. In this study, we aimed to answer the following research questions:

1. What effect did INSIGHTS have on changes in critical-thinking, language arts, and math skills of low-income shy children across kindergarten and the fall of first grade?

2. To what extent were intervention effects on critical-thinking and math skills mediated by behavioral engagement for shy students?

METHOD

Participants and Setting

Twenty-two elementary schools were partners in conducting this study. All schools served families with comparable sociodemographic characteristics in low-income neighborhoods. The participants included 345 children and their parents, as well as 122 teachers, from kindergarten and first-grade classrooms. Eleven of the schools were randomly assigned to host the INSIGHTS program, and the re-

maining 11 schools participated in the attention-control condition. The children included in this analysis were in kindergarten when they were enrolled in the study. They ranged in age from 4 to 7 years at baseline ($M = 5.38$ years, $SD = 0.61$ years). Half (50%) of the children were male. Eighty-seven percent of the children qualified for free or reduced-price lunch programs. Approximately 72% of children were Black and 19% were Hispanic non-Black, and the remaining children were biracial. A majority of the respondents were biological mothers (84%); other respondents included fathers (8%) and kinship guardians (7%). Approximately 28% of adult respondents did not finish high school, 26% had a high school or General Educational Development diploma, 24% had attended some college, and 22% had completed college.

Children enrolled in the study were similar in demographic characteristics to the other students at the schools who were invited but did not participate. According to school records, approximately 90% of the children in the partnering schools qualified for free or reduced-price lunch programs, and 78% of children were Black, 43% were Hispanic/Latino, 1% were White, and 6% were other races/ethnicities. School-level demographic characteristics were based on census categories in which race and ethnicity were not mutually exclusive categories.

Teacher participants included 60 kindergarten and 62 first-grade teachers (96% of whom were women). Sixty-one percent of the teachers reported their race/ethnicity as Black non-Hispanic, 10% as Hispanic/Latino non-Black, 23% as White, and 6% as Asian or biracial. All teachers had at least a bachelor's degree; 96% had a master's degree.

Recruitment and Randomization Procedures

Recruitment for this study was conducted by a racially and ethnically diverse team of field staff. All recruitment strategies were approved by university and school system research boards. Principals serving low-income students in three urban school districts were the first to be contacted. Team members explained the purpose of the study and its related logistics including randomization, after baseline data were collected, to one of two intervention conditions: INSIGHTS or a supplemental reading program. Twenty-three principals agreed to participate over 3 consecutive years, but one school withdrew from the study before randomization because of a transition in principals.

Kindergarten and first-grade teachers at participating schools were recruited individually or in small group meetings. Ninety-six percent of the kindergarten and first-grade teachers consented to participate. There was no attrition among the teachers for the duration of the study.

Parents from the participating kindergarten teachers' classrooms were recruited during September and October. Written materials describing the study and its procedures were sent home with the students. Interested parents were asked to contact the team. Parents were also recruited when they were at the school (e.g., during conference days or before or after school). After a parent consented, child assent was acquired.

All children enrolled in kindergarten at baseline in the participating schools were invited to take part in the study. However, given resource limitations and concern about teacher burden regarding data collection, no more than 10 representative students were recruited per classroom. Team members enrolled 4 to 10 students per classroom, or approximately 27% of the children attending kindergarten at the targeted schools. According to χ^2 tests, there were no significant differences between the children enrolled in the study and the children in the school as a whole in terms of percent female, Black, Hispanic, and eligible for free or reduced-price school lunch. After baseline data were collected, we used a random-numbers table to randomize schools to INSIGHTS or the supplemental reading program. Approximately half of the children and parents in the study participated in INSIGHTS ($n = 183$); the remaining child participants ($n = 162$) were enrolled in the attention-control condition. Similarly, approximately half of the

teachers ($n = 57$) participated in INSIGHTS; the remaining teachers ($n = 65$) were enrolled in the attention-control condition.

Measures

Student Academic Competence

The Academic Competency Evaluation Scale (ACES; DiPerna & Elliott, 2000) measured three subscales of teacher perceptions of the children's academic skills and achievement-related behaviors—critical thinking, language arts, and mathematics. Teachers rated students' academic skills in comparison with the grade-level expectations at their school (1, far below; 3, grade level; 5, far above). The critical-thinking subscale includes nine items and assesses higher-order thinking skills. The scale asks teachers to report on how well students engage in reflection, analysis, synthesis, and investigation. The reading/writing subscale consists of 11 items about the skills necessary for generating and understanding written language, including reading comprehension and written communication. The mathematics subscale primarily reflects skills related to the use and application of numbers. Thus the eight items it includes reflect measurement, computation, and problem solving.

The ACES was standardized on a large, national sample of elementary school teachers and students. Reliability for the ACES has been demonstrated through (a) high α coefficients for all grade clusters, (b) good test–retest correlations, and (c) adequate inter-rater correlations when rated by two different teachers (Demaray & Jenkins, 2011). In a study of 192 kindergarten through second-grade students, DiPerna, Volpe, and Elliott (2005) found high internal-consistency coefficients for the academic skills constructs, noting that they ranged from 0.92 to 0.98. In addition, 6-week test–retest stability coefficients for the scales were adequate, ranging from 0.81 to 0.92. In the current study, the average internal consistency across the three time points was 0.97 (critical thinking and reading/language arts) and 0.98 (math).

Student Behavioral Engagement

The Behavioral Observation of Students in Schools (Shapiro, 2004) was used to assess the frequency of students' behavioral engagement in academic activities. Each student was observed during two 15-min intervals on separate days. All observations occurred during morning academic activities involving literacy or mathematics instruction. Momentary time sampling procedures were used to measure the presence or absence of active engaged time (e.g., raising hand, writing) and passive engaged time (e.g., listening to instruction, reading silently) every 15 seconds. Each observation period was divided into 60 intervals (15 s each), in which a target student enrolled in the study was observed for four consecutive intervals, and a randomly selected peer was observed every fifth interval. This process was repeated across the 60 intervals, yielding 48 intervals of target student observation and 12 intervals of observation of randomly selected peers. Given that the study focused on the behavioral engagement of intervention students, only 96 observations of target students across two 15-min observation periods were used.

Child Temperament

The School-Aged Temperament Inventory (SATI; McClowry, 1995, 2002) was used to measure child temperament. The SATI is a 38-item 5-point Likert-type scale (ranging from *never* to *always*) that was standardized with a racially/ethnically and socioeconomically diverse sample of 883 parents reporting on their children. The instrument has four dimensions derived from principal-factor analysis: negative reactivity (12 items; intensity and frequency with which the child expresses negative affect), task persistence (11 items; degree of self-direction that a child exhibits in fulfilling task responsibilities), withdrawal (9 items; child's initial response to new people and situations), and activity (6 items; large motor activity). Examples of negative reactivity items include "gets upset when he/she can't find something" and "moody when corrected for misbehavior." Examples of task persistence are "returns to responsibilities after

friends call or visit" and "stays with homework until finished." Examples of items for withdrawal include "avoids new guests or visitors in the home," "seems nervous or anxious in new situations," and "prefers to play with someone he/she already knows than someone new." The activity dimension is composed of items like "runs to get where he/she wants to go" and "seems to be in a big hurry most of the time." In the current study, Cronbach's α values for the SATI were as follows: activity, $\alpha = 0.77$; withdrawal, $\alpha = 0.81$; task persistence, $\alpha = 0.85$; and negative reactivity, $\alpha = 0.87$. The α values were similar to those identified by McClowry (2002).

Data Collection

Researchers and field staff were provided group training on all procedures and measures before each of the three data collection periods. Time 1 data were collected at baseline in the winter (January/February) of the kindergarten year before the 10 weeks of intervention. Time 2 data were collected after the intervention in late spring of the kindergarten year (May/June). Time 3 data were collected in the fall (October/November) of first grade. Kindergarten teachers reported on children at Time 1 and Time 2, and first-grade teachers reported on children at Time 3. Data collection was scheduled at these times in the year to ensure that teachers had sufficient time to evaluate their students. Parents completed the SATI at baseline at their child's school via audio-enhanced computer-assisted self-interviewing software. Parents received $20 for their time. Teachers completed the ACES for each consented student and received $50 gift cards to purchase classroom supplies.

A team of data collectors was trained through (a) a 4-hr laboratory-based training session at the beginning of each year of the study (2008–2011), (b) three segments of video practice coding, (c) a 2-hr live training session in an elementary classroom, and (d) the achievement of 80% agreement or greater on all coding categories across two live classroom observations with a master coder. Data collectors conducted observations of behavioral engagement with the Behavioral Observation of Students in Schools at Times 1, 2, and 3. Because of resource constraints, only one data collector conducted observations in each classroom at a given time. Between the fall and spring observations within the academic year, data collectors were required to pass a continuing reliability test. Live reliability tests were conducted in the field. Initial and continuing reliability tests yielded inter-rater agreement that ranged from 0.80 to 0.95.

Intervention Procedures

Facilitator Training

Before conducting the intervention in the schools, INSIGHTS facilitators attended a fall-semester graduate-level course, which met once a week for approximately 2 hr, to learn the underlying theory and related research. New facilitators were also trained by experienced facilitators to use the intervention materials. Facilitators were graduate students in psychology, education, and educational theater with varied racial/ethnic backgrounds. A total of eight facilitators over the course of the study were trained to deliver INSIGHTS. Each facilitator conducted the full intervention (teacher, parent, and child/classroom) in the school to which he or she was assigned.

Program Delivery

Beginning in January, teachers and parents attended 10 two-hour, facilitated sessions based on a structured curriculum that included didactic content and professionally produced vignettes as well as handouts and group activities. One session was attended by parents and teachers together; the others were conducted separately. Makeup sessions were offered as needed. For each session attended, parents received $20 and teachers received $40 gift cards. The teachers also received professional development credit.

The classroom program was delivered in 45-min lessons during the same 10 weeks as the parent and teacher sessions. The curriculum materials included puppets, workbooks, flash cards, and videotaped vignettes. Although the facilitators had primary responsibility for conducting the classroom sessions,

teachers participated in the sessions, especially when the students practiced resolving dilemmas. No makeup sessions were conducted, although teachers were asked to use the program materials with the students during the week.

Attendance

The average number of teacher sessions attended was 9.44 ($SD = 0.91$). The majority of teachers attended all sessions (70.6%), and another 26.5% attended eight or nine sessions. The average number of classroom sessions attended by the participating children was 8.30 ($SD = 2.25$). Thirty-two percent of children were present for all classroom sessions, and 46.3% were present for eight or nine sessions. The average number of sessions attended by parents of participating children was 5.93 ($SD = 4.15$). Twenty-five percent of the parents were present for all sessions, and 30.3% were present for eight or nine sessions. This amount of intervention dosage is comparable with similar socioemotional learning interventions (e.g., 4Rs: Brown, Jones, LaRusso, & Aber, 2010; Chicago School Readiness Project: Raver et al., 2011; Incredible Years: Webster-Stratton, Reid, & Stoolmiller, 2008).

Fidelity

To maintain model fidelity, facilitators followed scripts, used material checklists, and documented sessions. Deviations or clinical concerns were discussed weekly in supervision with the program developer. Supervision focused on challenges related to conducting sessions, implementation logistics, and participant concerns. All teacher and parent sessions were videotaped and reviewed for coverage of content and effectiveness of facilitation (Hulleman & Cordray, 2009). Videotapes showed that 94% of the curriculum was adequately covered in the teacher sessions and 92% of the curriculum was covered in the parent sessions.

Attention-Control Condition

Schools not randomized to INSIGHTS hosted a supplemental reading program as an attention-control condition to ensure that all children in the participating schools were provided with additional support. Students in attention-control schools participated in a 10-week after-school reading program. In addition, their teachers attended two workshops focused on strategies to enhance early literacy. Similar content was provided in the two workshops for parents. There was no overlapping content between the supplemental reading program and INSIGHTS.

The average number of child sessions attended in the supplemental reading program was 8.14 ($SD = 2.26$). Thirty-four percent of children participated in 10 sessions; an additional 38% took part in 8 or 9 sessions. Thirty percent of parents and 83% of teachers attended both workshops. Sixty-two percent of parents and 17% of teachers attended one. No makeup sessions were provided. Parents received $20 and teachers received $40 for classroom resources for each workshop. The participants also received reading materials.

Analytic Approach

Missing-Data Analysis

Child study participants who had at least two time points of data were included in the current analysis. After beginning with a total of 374 children, 29 were excluded from the analysis because they had only one data point. We used independent-samples t tests for continuous variables (parental education, child temperament) and χ^2 tests for categorical variables (race, gender, free lunch eligibility, parental marital status) to determine that children excluded from analyses did not differ from children included. For the remaining children, missingness was low to moderate for our predictors and outcomes, ranging from 0% to 21% for time-varying and time-invariant variables. To include all children who remained in the study at the third time point, missing values for continuous variables were imputed using a Markov chain Monte Carlo method (Schafer, 1997) in SAS PROC MI. This was appropriate because the data were normally distributed (Graham & Donaldson, 1993; Kellam, Rebok, Ialongo, & Mayer, 1994). The Markov chain Monte Carlo method uses sim-

ulation from a Bayesian prediction distribution. Ten imputations were performed with a burn-in period of 500. Interaction terms were also created before imputation and were imputed separately. All conditional analyses were run 10 separate times in HLM 7 (Raudenbush, Bryk, Cheong, & Congdon, 2011), and final parameter estimates were generated by calculating the mean of these 10 estimates.

Growth-Curve Modeling

Because data on students' critical-thinking, language arts, and math skills were collected before the intervention in kindergarten and at two subsequent time points (the spring of kindergarten and the fall of first grade), individual growth modeling was used to examine change over time in each of these skills. Individual growth modeling allows one to model change over time in an outcome with repeated measures (Singer & Willett, 2003). All models were fitted with HLM 7, using a maximum likelihood estimator (Raudenbush et al., 2011). The metric of time used was assessment point. Time was centered at Assessment Point 3 so that the parameter for the intercept would represent the outcomes at the final intervention follow-up point. To center time, the number three was subtracted from the time (assessment point) metric.

To accurately estimate contextual school-level effects on students' critical-thinking, language arts, and math skills, continuous predictors at Levels 1 and 2 were centered around their group mean. Group mean centering produces a more accurate estimate of the Level 1 (individual) slope; this, in turn, allows for a more precise estimation of the moderating effect of Level 2 (student) predictors on time-varying student outcomes (Enders & Tofighi, 2007; Raudenbush & Bryk, 2002). Level 2 categorical variables (female, child Black, child Hispanic) were not centered. Preliminary analyses indicated that the Level 3 (school) predictors for school size, percent free/reduced-price lunch, percent Black, and percent Hispanic were bimodal. As such, these were included as dichotomous variables (1, above the sample mean; 0, below the sample mean) in all models.

Initial analyses consisting of unconditional models were run for each of the children's academic skills (critical thinking, language arts, and math) to determine whether there was significant between-individual and between-school variation in these predictors. Then, to address the substantive questions proposed in this study, four sets of models were tested. First, an unconditional baseline model with no predictors was run for math, language arts, and critical thinking to determine the proportion of variance attributed to individual- and school-level factors. On the basis of the estimates obtained from the unconditional model, intraclass correlations were computed. Intraclass correlations represent the proportion of total variance attributed to mean differences between individuals and schools. Unconditional models suggested there was significant between-individual and between-school variation in these data. As such, a random effect was included at Level 2 and Level 3 in all models, allowing the intercept to vary for these two levels of nesting (Raudenbush, 2009). The equation for the unconditional-means model is as follows:

$$\text{Outcome}_{tij} = \gamma_{00} + u_{0ij} + \upsilon_{0j} + r_{tij} \quad (1)$$

The subscript t refers to repeated response variable observations (Level 1 units) collected from i children (Level 2 units) over time (Peugh, 2010) in School j. The model in Equation 1 is called an unconditional-means model because the academic competence scores for Student i at Time t are modeled as a function of (a) a grand mean academic competence score for all children (γ_{00}), (b) a term that represents deviations in an individual's academic competence mean around the grand academic competence mean (u_{0ij}), (c) a term that represents deviations in the school academic competence (υ_{0j}), and (d) a time-specific residual term that demonstrates the differences between each individual's observed academic skill competence and predicted academic skill competence (r_{tij}; Peugh, 2010).

Second, an unconditional-growth model was fitted to examine children's academic skill

scores from the first assessment point through Assessment Point 3.

$$\text{Outcome}_{ij} = \gamma_{00} + \gamma_{10}\,(\text{Assessment point}_{ij}) + u_{0ij} + \upsilon_{0j} + r_{tij} \quad (2)$$

As shown in Equation 2, each student's academic competence score at the intercept is modeled as a grand mean academic competence score at Assessment Point 1 (γ_{00}), as well as a residual term that demonstrates deviations in academic competence scores at the first assessment point about the grand mean (u_{0ij}). In addition, each student's rate of academic competence score change across time is modeled as a grand mean rate of academic competence change (γ_{10}).

Third, a conditional model (Model 1) was run in which the Level 2 predictors for shyness and treatment were entered into the model. At the same time, a series of Level 2 covariates—(a) child gender (binary; boys: 0, girls: 1), (b) child Black (binary; no: 0, yes: 1), (c) child Hispanic (binary; no: 0, yes: 1), (d) negative reactivity (continuous: 1–5), (e) task persistence (continuous: 1–5), and (f) activity (continuous: 1–5)—were added as Level 2 time-invariant predictors to account for between-child variation in academic competence within schools. Finally, a series of Level 3 covariates—(a) study cohort (dummy codes were included for Cohort 1 and Cohort 2, whereas Cohort 3 was the referent group; estimates from the cohort covariate are not included in the tables for ease of presentation), (b) school size, (c) school percent free/reduced-price lunch, (d) school percent Black, and (e) school percent Hispanic—were added to account for between-school variation in academic competence. We controlled for these preintervention family, child, and school characteristics to increase the power of the randomized study design when examining the influences of a shy temperament and INSIGHTS on children's academic skills.

A final model (Model 2) was run in which cross-level interactions between time (Level 1), treatment (Level 2), and shyness (Level 3) were entered into the existing model. Significant cross-level interactions indicate that the time-varying relationship between time and shyness varies as a function of treatment. To identify a precise estimate of the growth effect of treatment for shy children over and above the influence of other growth trends by temperament, we also included cross-level interactions between time, treatment, and the remaining dimensions of temperament (negative reactivity, task persistence, and activity) in this step. The main effects highlight the within-time effect of treatment and shyness on academic competence, controlling for the remaining dimensions of temperament. Significant interaction terms indicate differential growth in academic competence over time for shy students in the treatment group compared with shy students in the attention-control condition. Effect sizes for statistically significant findings were calculated following procedures described by Feingold (2009) for growth model analysis, yielding effect sizes in the same metric as classical designs, thus facilitating comparisons across studies.

In cases in which we observed significant effects of the INSIGHTS × Shyness × Time cross-level interaction, we examined the mediating role of behavioral engagement in explaining the effects of INSIGHTS for shy children on growth in academic skills. As a preliminary step, we first included behavioral engagement, predicting both the intercept and the slope, in the fully interacted model (Model 2) to test whether the coefficient for any significant interaction terms predicting the slope decreased with the addition of behavioral engagement.

In these planned analyses, treatment was a Level 2 (time-invariant individual) variable and the mediator (behavioral engagement) and outcomes (academic competence) were Level 1 (time-varying individual) variables. We used a mediation framework developed by Zhang, Zyphur, and Preacher (2009), using steps similar to those developed by Baron and Kenny (1986). The first step of the mediation analysis involved assessment of the effects of treatment condition on the outcomes (academic competence in critical thinking, language arts, and math) controlling for Level 2

and Level 3 covariates (Path C). In this framework, Path C represents the direct relationship between treatment and the outcome. In the second step, we assessed the effects of treatment on the mediator (behavioral engagement, Path A). In the third and final step of the mediation analysis, we assessed the effects of treatment condition and the Level 2 group mean of the mediator (behavioral engagement) on the outcome (academic competence), controlling for child and school covariates (Paths B and C'). Path C' considers the relationship between treatment and the outcome, taking into account the effect of the mediator (Zhang et al., 2009). In this step, we were primarily interested in examining whether the coefficient for any of the statistically significant interaction terms from Model 2 (INSIGHTS × Shyness × Time) decreased with the addition of behavioral engagement as a predictor.

In the case that we did observe initial evidence for mediation in the full sample, we then planned to test the mediated paths within the subset of shy children to improve the interpretability of the findings for this subgroup of interest. In previous research, McClowry (2002) identified cut points for students designated to be "high in withdrawal" based on students scoring a 2.8 on this dimension of the SATI. In this study, we also operationalized shyness as children scoring more than half of a standard deviation above the mean level on the shyness variable. Because the results of these analyses were similar, we report the findings from the more conservative approach using the standard deviation cut score. As such, children who scored greater than 3.02 on the shyness scale were included in the mediation analyses. There were 50 INSIGHTS children in the shy group (mean = 3.67, $SD = 0.49$, minimum = 3.10, maximum = 5) and 53 children in the comparison attention-control shy group (mean = 3.65, $SD = 0.40$, minimum = 3.03, maximum = 4.72). We used independent-samples *t* tests to determine whether there were significant baseline differences between the children in the INSIGHTS condition and those in the attention-control condition within the shy mediation analysis subgroup. We found no significant differences between treatment and attention control for critical-thinking, math, or language skills; behavioral engagement; and shyness levels within this subgroup. In addition, we did not identify any significant pretreatment differences between the shy and nonshy children for the three remaining dimensions of temperament. Thus, we continued with our mediation models using the procedure outlined above (Zhang et al., 2009).

RESULTS

Means and standard deviations for continuous variables and percentages for dichotomous variables (by treatment/control) are presented in Table 1. Independent-samples *t* tests showed no significant pretreatment differences between children enrolled in INSIGHTS and children enrolled in the supplemental reading program with respect to the observed continuous predictor and outcome variables, including the four dimensions of temperament, used in these analyses. In addition, χ^2 analyses to examine differences between participants enrolled in the INSIGHTS treatment and those enrolled in the attention-control condition suggested no significant differences in the groups by child Black, gender, or free lunch eligibility (see Table 1). However, χ^2 analyses did suggest that there were more Hispanic children enrolled in the INSIGHTS condition relative to the attention-control condition.

Change in Academic Skills

The results of the unconditional-means model shown in Equation 1 showed significant grand mean scores for the three dimensions of academic skills (critical thinking: $b = 2.72$, $p < .01$; language arts: $b = 2.65$, $p < .01$; math: $b = 2.70$, $p < .01$). The results also showed that children's mean scores for academic skills (i.e., the mean score across all assessments) significantly varied around the grand mean. Intraclass correlation calculations indicated that 33.78% of the variation in critical-thinking skills, 48.79% of the variation in language arts skills, and 41.24% of the variation in math skills occurred across students.

Table 1. Means and Standard Deviations for Study Variables

	Time 1				Time 2				Time 3			
	Treatment		Control		Treatment		Control		Treatment		Control	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
ACES for reading	2.64	0.76	2.71	0.68	2.80	0.72	2.89	0.69	2.49	0.89	2.46	0.80
ACES for math	2.67	0.66	2.72	0.58	2.79	0.65	2.91	0.58	2.52	0.73	2.63	0.67
ACES for critical thinking	2.65	0.64	2.72	0.60	2.82	0.62	2.95	0.58	2.56	0.65	2.58	0.65
Behavioral engagement	0.68	0.18	0.72	0.17	0.69	0.26	0.66	0.22	0.69	0.15	0.73	0.15
Child female	0.48	0.50	0.52	0.51								
Child Black	0.74	0.43	0.70	0.46								
Child Hispanic	0.22	0.40	0.15	0.36								
Shyness	2.51	0.88	2.65	0.88								
Negative reactivity	2.85	0.88	2.93	0.83								
Task persistence	3.89	0.82	3.69	0.97								
Activity	2.88	0.85	2.92	0.88								
School size	469.09	132.51	467.73	202.01								
% eligible free/reduced lunch	90	32	90	32								
% Black	82	9	74	17								
% Hispanic	49	28	37	27								

Note. An independent-samples *t* test yielded a significant pretest difference ($p < .05$) between the treatment and control groups for child Hispanic. The sample size was 345 children.
Abbreviation: ACES, Academic Competency Evaluation Scale.

The results for the unconditional-growth model in Equation 2 showed significant grand mean academic skill scores at Assessment Point 3 for critical-thinking ($b = 2.68$, $p < .01$), language arts ($b = 2.54$, $p < .01$), and math ($b = 2.62$, $p < .01$) scores. The findings indicated that scores decreased on average 0.04 ($p < .01$) at each time point for critical thinking, 0.11 ($p < .01$) for language arts, and 0.07 ($p < .01$) for math. Furthermore, variance component estimates showed (a) significant variance in observed versus predicted academic skill scores within students (Level 1 residual critical thinking: $\tau_{10} = 0.13$, $p < .01$; Level 1 residual language arts: $\tau_{10} = 0.32$, $p < .01$; Level 1 residual math: $\tau_{10} = 0.20$, $p < .01$) and (b) significant slope variance (critical thinking: $\tau_{11} = 0.04$, $p < .05$; language arts: $\tau_{11} = 0.03$, $p < .05$; math: $\tau_{11} = 0.06$, $p < .01$) in academic skill trajectories across students.

Model 1

Given that the unconditional-growth model showed significant intercept and slope variance in critical-thinking, language arts, and math scores across children, predictor variables were added to Level 1, Level 2, and Level 3 of the model to explain this variance. The values presented for Model 1 in Table 2 indicate the association between the independent variables and critical thinking, language arts, and math after controlling for the other effects in the model and can be interpreted as partial correlations. Several significant predictors of the intercept were found. Specifically, children with shyer temperaments evidenced lower scores on teacher reports of critical-thinking ($b = -0.09$, $p < 01$), language arts ($b = -0.16$, $p < .01$), and math ($b = -0.09$, $p < 01$) skills than their less shy peers. A

Table 2. Model Summary for Growth Models Examining Academic Skill Scores

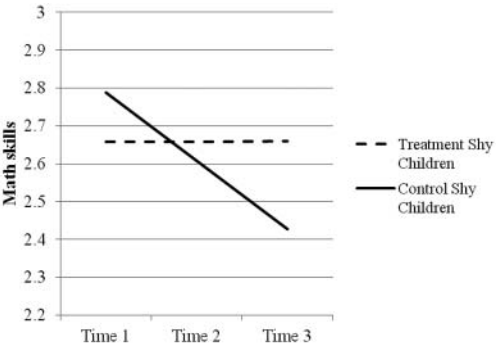
	Model 1						Model 2					
	Critical Thinking		Language Arts		Math		Critical Thinking		Language Arts		Math	
Model Parameter	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Intercept	2.81*	0.15	2.74*	0.17	2.87*	0.19	2.83*	0.15	2.75*	0.23	2.88*	0.19
Time	–0.04*	0.02	–0.11*	0.03	–0.07*	0.02	–0.03*	0.03	–0.10*	0.03	–0.06*	0.03
Treatment	–0.06	0.07	–0.10	0.13	–0.07	0.09	–0.08	0.07	–0.12	0.12	–0.10	0.10
Child female	0.11*	0.05	0.19*	0.09	0.08	0.06	0.10*	0.05	0.18*	0.07	0.07	0.06
Child Black	–0.14*	0.07	–0.18*	0.10	–0.19*	0.08	–0.15*	0.07	–0.18*	0.10	–0.19*	0.08
Child Hispanic	–0.27*	0.08	–0.32*	0.10	–0.31*	0.09	–0.27*	0.08	–0.32*	0.11	–0.32*	0.09
School size	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% eligible for free/ reduced-price lunch	–0.03	0.13	–0.06	0.14	–0.09	0.17	–0.03	0.13	–0.05	0.21	–0.09	0.17
Percent Black	0.01	0.33	0.28	0.36	0.28	0.43	0.01	0.33	0.28	0.53	0.28	0.43
Percent Hispanic	0.03	0.16	0.11	0.30	0.07	0.21	0.02	0.16	0.11	0.26	0.06	0.21
Shyness	–0.09*	0.03	–0.14*	0.04	–0.09*	0.03	–0.09*	0.04	–0.16*	0.05	–0.04	0.03
Negative reactivity	0.03	0.04	0.02*	0.04	–0.02	0.04	0.07*	0.04	0.05	0.06	0.01	0.04
Task persistence	0.08*	0.03	0.10*	0.04	0.09*	0.03	0.07*	0.04	0.09*	0.05	–0.02	0.03
Activity	0.01	0.03	0.03	0.03	0.04	0.04	–0.01	0.04	–0.01	0.05	–0.04	0.03
Treatment × Time							–0.02	0.04	–0.03	0.04	–0.03	0.04
Shyness × Time							–0.03	0.03	–0.05	0.03	–0.04	0.03
Negative reactivity × Time							0.04	0.04	0.02	0.04	0.01	0.04
Task persistence × Time							0.00	0.03	–0.03	0.04	–0.02	0.03
Activity × Time							–0.04	0.03	–0.05	0.04	–0.04	0.03
Treatment × Shyness × Time							0.08*	0.04	0.06	0.04	0.07*	0.03
Treatment × Negative reactivity × Time							–0.02	0.04	0.01	0.05	–0.06	0.05
Treatment × Task persistence × Time							–0.03	0.03	–0.03	0.04	–0.02	0.03
Treatment × Activity × Time							0.05	0.04	0.02	0.05	0.07*	0.04
Deviance statistic	1,781.36		2,096.94		1,862.67		1,770.13		2,090.12		1,876.38	

Note. The sample size was 345 children.
*$p < .05$.

significant main treatment effect was not detected for any of the outcomes.

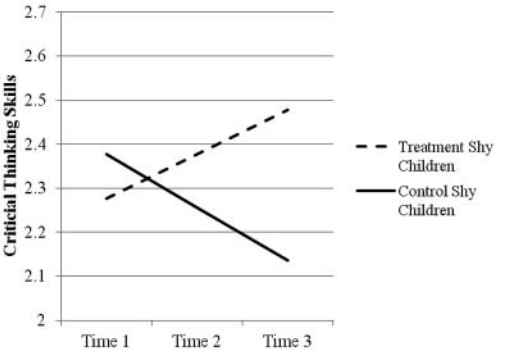
Variance component estimates showed (a) significant variance in observed versus predicted critical-thinking, language arts, and math scores within students (critical thinking, Level 1 residual: $\sigma^2 = 0.24$, $p < .01$; language arts, Level 1 residual: $\sigma^2 = 0.27$, $p < .01$; math, Level 1 residual: $\sigma^2 = 0.24$, $p < .01$) and (b) significant variation in critical-think-

Figure 3. Growth Model Predicting Math Skills of Shy Children



Note. Estimated effects include all covariates. Mean differences in the outcome at Time 1 (baseline) were not statistically significant.

Figure 4. Growth Model Predicting Critical-Thinking Skills of Shy Children



Note. Estimated effects include all covariates. Mean differences in the outcome at Time 1 (baseline) were not statistically significant.

ing, language arts, and math scores at Assessment Point 3 (critical thinking: $\tau_{00} = 0.11$, $p < .01$; language arts: $\tau_{00} = 0.27$, $p < .01$; math: $\tau_{00} = 0.17$, $p < .01$). The Level 1 residual variance and the Level 2 intercept and slope variance estimates decreased for all outcomes, indicating that the independent variables in the model were relatively strong predictors of critical-thinking, language arts, and math skills within and between individuals.

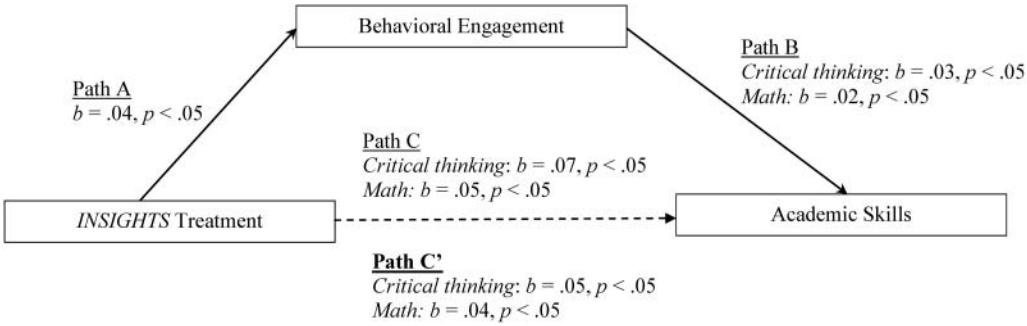
Model 2

To test whether the effect of the time-varying relationship between time and shyness varied as a function of treatment, a series of cross-level interactions were added to the model. A significant effect of the three-way interaction, Treatment × Time × Shy, was found for both critical-thinking skills ($b = 0.08$, $p < .05$, ES = 0.40) and math ($b = 0.07$, $p < .05$, ES = 0.36). As illustrated by Figure 3, shy children in the treatment group experienced stable math skill scores across the transition to elementary school relative to shy children in the attention-control condition, who showed declines in math skills. Similarly, as shown in Figure 4, the results showed that shy children in the treatment group improved in their critical-thinking skills across this transition, as compared with the children in the attention-control condition, who declined in critical-thinking skills. In addition, Table 2 shows that the Treatment × Time × Activity interaction was statistically significant for math achievement ($b = 0.07$, $p < .05$). However, because this study is focused on understanding intervention effects for shy children, we will not discuss this finding in depth.

Mediation Effects

Multilevel mediation analyses were conducted to examine whether intervention effects on critical-thinking and math skills were mediated by behavioral engagement for shy students. We initially examined the mediating effect of behavioral engagement on academic competence for the full sample of students. We found a significant, direct effect of behavioral engagement on critical-thinking ($b = 0.05$, $p < .05$, ES = 0.31) and math ($b = 0.04$, $p < .04$, ES = 0.25) skills but not language arts skills. Of note, after accounting for the effect of behavioral engagement, the *b* coefficient for the interaction between shyness, treatment, and time decreased in the models predicting critical-thinking ($b = 0.05$, $p < .05$) and math ($b = 0.04$, $p < .05$) skills. This finding provides initial evidence that the effect of treatment on shy children's critical-thinking and math skills was partially ex-

Figure 5. INSIGHTS and Behavioral Engagement Predicting Academic Skills



Note. Path C refers to the direct effect of INSIGHTS on academic skills, whereas Path C' refers to the effect of INSIGHTS on academic skills controlling for the mediator behavioral engagement. The behavioral engagement mediator predicted in Path A is a Level 1 (time-varying) measure, whereas the behavioral engagement mediator predicting academic skills in Path B is a Level 2 (time-invariant) group mean measure, per instructions for modeling multilevel mediation recommended by Zhang et al. (2009).

plained by an improvement in behavioral engagement.

In the second part of the analysis, we examined this mediating effect of behavioral engagement on dimensions of academic competence within the subsample of shy children. As exhibited in Figure 5, shy children in INSIGHTS exhibited faster growth in critical thinking and math over the three time points of the study, as compared with shy children in the attention-control condition (critical thinking, Path C: $b = 0.07$, $p < .05$; math, Path C: $b = 0.05$, $p < .05$). Treatment also predicted growth in behavioral engagement for shy children (Path A: $b = 0.04$, $p < .05$). A final model showed that the effect of treatment on growth in critical thinking and math was partially mediated through growth in behavioral engagement (critical thinking, Path B: $b = 0.03$, $p < .05$: math, Path B: $b = 0.02$, $p < .05$; critical thinking, Path C': $b = 0.05$, $p < .05$; math, Path C': $b = 0.04$, $p < .05$). An examination of the b coefficients and their associated significance levels indicated that the positive effects of intervention type were partially mediated through gains in behavioral engagement for shy children in the treatment group. These findings show that for shy children, there were direct effects of the treatment on the math and critical-thinking skill outcomes, as well as indirect effects of the treatment on math and critical thinking, occurring through behavioral engagement.

DISCUSSION

A shy temperament is an identified risk factor for compromised academic skills (Evans, 2010), but few studies have explored the developmental trajectories of shy children's academic skills. Even fewer studies have examined the effects of interventions on shy children's academic development and the mechanisms through which interventions support that development. This study investigated the impact of a temperament-based intervention, INSIGHTS, in enhancing the academic development of shy children across kindergarten and at the beginning of first grade. The participating children in the study were from low-income families and they attended under-resourced urban schools.

Effects of INSIGHTS on Academic Skills

Children in INSIGHTS and in the attention-control condition showed gains in math, language, and critical-thinking skills during kindergarten. However, consistent with previous studies (Evans, 2010), we found that shyer children evidenced lower-level academic skills than their less shy peers. More specifi-

cally, children with shyer temperaments evidenced lower scores on teacher reports of critical-thinking, language arts, and math skills than their less shy peers. However, differential intervention effects on shy children were found. Shy children in INSIGHTS evidenced statistically significant growth in critical-thinking skills and stability in math skills over the transition from kindergarten to first grade compared with their shy peers in the attention-control condition, who declined in both domains. These findings are of note because national data show that students' critical thinking on standardized tests is as important as the acquisition of basic skills and facts (Wenglinsky, 2004). Likewise, a study by Duncan et al. (2007) found that early math achievement was the single most powerful predictor of future educational attainment. Children who persistently score in the bottom end of the math distribution in elementary school are 13 percentage points less likely to graduate from high school and 29 percentage points less likely to attend college (Duncan, 2001).

Although there were positive outcomes for critical thinking and math, there were no gains in language arts skills for shy children attributed to INSIGHTS. The control condition was a supplemental after-school reading program, which might have contributed to the finding. The shy children in the supplemental reading program may have benefited from its language arts content and the reading coaches' use of a small-group format.

The Role of Behavioral Engagement

Theorists have proposed that behavioral engagement is a mediating mechanism through which shyness compromises the acquisition of academic skills (Hughes & Coplan, 2010). The demands of the school environment can exacerbate shy children's social reticence and self-consciousness and thereby inhibit their behavioral engagement (Evans, 2001). Moreover, teachers who perceive children as disengaged tend to rate these children's academic skills as lower than those of their peers regardless of their actual performance (Hughes & Coplan, 2010). In fact, some studies have found few differences between shy children and their nonshy peers in performance on standardized assessments of academic skill development yet have found significant differences in teacher reports (Hughes & Coplan, 2010).

Unlike temperament, behavioral engagement is malleable, responsive to contextual features, and amenable to environmental change during childhood (Fredericks, Blumenfeld, & Paris, 2004). Our findings suggest that the effects of INSIGHTS on shy children's critical thinking and math may be partly explained by changes in their behavioral engagement. Shy children in INSIGHTS evidenced more rapid increases in behavioral engagement than their nonshy peers. These higher rates of engagement were, in turn, linked to better math and critical-thinking skills for shy children in INSIGHTS compared with those in the attention-control condition. The increase in engagement among young shy children in INSIGHTS is important. Disengaged children in the primary grades do not catch up to their peers as they progress through elementary school (Angus et al., 2009). Instead, student engagement in first grade predicts both standardized achievement test scores and teacher-assigned grades throughout elementary school (Alexander, Entwisle, & Horsey, 1997), as well as academic attainment by early adulthood (Entwisle, Alexander, & Olson, 2005).

Engagement only partially mediated the pathway between intervention effects and critical thinking and math because about 30% of the effect of INSIGHTS on critical thinking and 20% of the effect of INSIGHTS on math were explained by behavioral engagement. Other mechanisms may have contributed to the links between shyness and academic skill development. For example, shyness in elementary school is concurrently and predictively associated with peer relationship difficulties (e.g., exclusion, victimization) and internalizing problems such as anxiety, loneliness, and low self-esteem (Coplan, Arbeau, & Armer, 2008; Coplan, Closson, & Arbeau, 2007; Gazelle & Ladd, 2003; Hart et al., 2000).

Limitations and Future Research

This study had several limitations that need to be addressed, which provide directions for future research. First, we only examined intervention effects on academic performance in kindergarten and at the start of first grade. Although the findings highlight the importance of intervention in kindergarten, future research should follow children across the early years of elementary school so that trajectories of change in academic skills can be investigated. Second, although we used observational, parent- and teacher-report measures of temperament, classroom behaviors, and academic skill development, the inclusion of additional measures would have been desirable, in particular standardized measures of student academic skills. Third, although the sample represents urban schools with high proportions of low-income students, which is a population prioritized for early intervention, the generalizability of the findings is limited. Consequently, we are unable to ascertain the intervention effects for children with different sociodemographic and racial/ethnic backgrounds.

Fourth, further study is needed to examine why INSIGHTS appears particularly effective with shy children and, to a lesser extent, with children high in activity (with respect to math performance) but is not associated with improved academic outcomes for the remaining two temperament styles (task persistence and negative reactivity). We were not especially surprised that the intervention proved particularly effective for shy children and, to a lesser extent, active children, given the task demands faced by children at the transitions to kindergarten and first grade. More specifically, at these key points, children are forced to assert themselves more in terms of academic participation and to attend for longer periods to academic tasks than in previous school experiences. Children with high task persistence do not struggle with these demands, and those higher in negative reactivity may also not find participating in class discussions or focusing on academic tasks to be challenging. As such, the intervention may not substantially improve the goodness of fit, regarding the academic environment, of the classroom environment for these children.

Implications for Practice and Social Policy

This study has implications for school personnel, including interventionists, school psychologists, and teachers. In the fast-paced world of classrooms, especially those with high concentrations of children with poverty-related stressors, teachers often focus on students with disruptive behavior problems. As a result, shy students often receive too little attention. However, the effects of INSIGHTS on shy children suggest that goodness of fit can enhance the classroom context to meet their particular socioemotional needs. INSIGHTS may have enabled teachers to recognize and appreciate these children's particular temperament style. By using the scaffold-and-stretch approach, teachers might also have helped their shy students achieve success in situations otherwise avoided. Likewise, it is possible that the shy children may have felt more accepted and supported by their peers. Perhaps the responsivity of the environment led to the shy children feeling more comfortable, thus supporting their behavioral engagement when asked to perform a cognitively challenging task such as solving a math problem or answering an open-ended question requiring critical thinking.

The current study shows that comprehensive interventions can enhance the academic skill development of urban children at risk for lower critical-thinking and math skills because of a shy temperament. No other intervention studies have improved the behavioral engagement of shy children as a means of supporting their academic development. Our findings indicate that changes in behavioral engagement are an important mechanism through which preventive interventions can influence their academic development. Moreover, the results presented here show that a preventive intervention focusing on enhancing teachers' awareness of and responsiveness to child temperament may be particularly helpful for children with shy temperaments.

REFERENCES

Alexander, K. L., Entwisle, D. R., & Horsey, C. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education, 70,* 87–107.

Alexander, K. L., Entwisle, D., & Kabbani, N. (2001). The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record, 103,* 760–822.

Angus, M., McDonald, T., Ormond., C., Rybarcyk, R., Taylor, A., & Winterton, A. (2009). *Trajectories of classroom behaviour and academic progress: A study of student engagement with learning.* Mount Lawley, Australia: Edith Cowan University.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Bossaert, G., Doumen, S., Buyse, E., & Verschueren, K. (2011). Predicting children's academic achievement after the transition to first grade: A two-year longitudinal study. *Journal of Applied Developmental Psychology, 32,* 47–57.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs program. *Journal of Educational Psychology, 102,* 153–167.

Bruce, J., Davis, E. P., & Gunnar, M. R. (2002). Individual differences in children's cortical response to the beginning of a new school year. *Psychoneuroendocrinology, 27,* 635–650.

Chess, S., & Thomas, A. (1984). *Origins and evolution of behavior disorders.* Cambridge, MA: Harvard University Press.

Coplan, R. J. (2000). Assessing nonsocial play in early childhood: Conceptual and methodological approaches. In K. Gitlin-Weiner, A. Sandgrund, & C. Schaefer (Eds.), *Play diagnosis and assessment* (2nd ed., pp. 563–598). New York: Wiley.

Coplan, R. J., Arbeau, K. A., & Armer, M. (2008). Don't fret, be supportive! Maternal characteristics linking child shyness to psychosocial and school adjustment in kindergarten. *Journal of Abnormal Child Psychology, 36,* 359–371.

Coplan, R. J., Closson, L. M., & Arbeau, K. A. (2007). Gender differences in the behavioral associates of loneliness and social dissatisfaction in kindergarten. *Journal of Child Psychology and Psychiatry, 48,* 988–995.

Coplan, R. J., Findlay, L. C., & Nelson, L. J. (2004). Characteristics of preschoolers with lower perceived competence. *Journal of Abnormal Child Psychology, 32,* 399–408.

Crozier, W. R., & Perkins, P. (2002). Shyness as a factor when assessing children. *Educational Psychology in Practice, 18,* 239–244.

Curby, T. W., Rimm-Kaufman, S., & Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology, 101,* 912–925.

Demaray, M. K., & Jenkins, L. N. (2011). Relations among academic enablers and academic achievement in children with and without high levels of parent-rated symptoms of inattention, impulsivity, and hyperactivity. *Psychology in the Schools, 48,* 573–586.

Denham, S. A., & Brown, C. (2010). "Plays nice with others": Social–emotional learning and academic success. *Early Education & Development, 21,* 652–680.

DiPerna, J. C., & Elliott, S. N. (2000). *Academic competence evaluation scales.* San Antonio, TX: Psychological Corporation.

DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in the elementary grades. *Journal of School Psychology, 43,* 379–392.

Dobbs, J., Doctoroff, G. L., Fisher, P. H., & Arnold, D. H. (2006). The association between preschool children's socio-emotional functioning and their mathematical skills. *Journal of Applied Developmental Psychology, 27,* 97–108.

Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning? *School Psychology Review, 36,* 413–432.

Duncan, G. (2001). *The importance of academic skills for preK–3rd. Report to the Foundation for Child Development.* Retrieved from http://fcd-us.org/resources/importance-academic-skills-prek-3rd

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428–1446.

Elias, M. J., & Haynes, N. M. (2008). Social competence, social support, and academic achievement in minority, low-income, urban elementary school children. *School Psychology Quarterly, 23,* 474–495.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12,* 121–138.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22, a new story. *American Journal of Sociology, 110,* 1458–1502.

Evans, M. (2010). Language performance, academic performance, and signs of shyness: A comprehensive review. In K. H. Rubin & R. J. Coplan (Eds.), *The development of shyness and social withdrawal.* New York: Guilford.

Evans, M. A. (1987). Discourse characteristics of reticent children. *Applied Psycholinguistics, 8,* 171–84.

Evans, M. A. (2001). Shyness in the classroom and home. In W. R. Crozier & L. E. Alden (Eds.), *International handbook of social anxiety: Concepts, research and interventions relating to the self and shyness* (pp. 159–183). Sussex, England: Wiley.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P. A., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review, 36,* 44–62.

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14,* 43–53.

Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74,* 59–109.

Gazelle, H., & Ladd, G. W. (2003). Anxious solitude and peer exclusion: A diathesis stress model of internalizing trajectories in childhood. *Child Development, 74*, 257–278.

Gersten, M. (1989). Behavioral inhibition in the classroom. In J. S. Reznick (Ed.), *Perspectives on behavioral inhibition* (pp. 71–91). Chicago: University of Chicago.

Gordon, E. M., & Thomas, A. (1967). Children's behavioral style and the teacher's appraisal of their intelligence. *Journal of School Psychology, 5*, 292–300.

Graham, J. W., & Donaldson, S. I. (1993). Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and use of follow-up data. *Journal of Applied Psychology, 78*, 119–128.

Hart, C. H., Yang, C., Nelson, L. J., Robinson, C. C., Olsen, J. A., Nelson, D. A., . . . Wu, P. (2000). Peer acceptance in early childhood and subtypes of socially withdrawn behavior in China, Russia and the United States. *International Journal of Behavioral Development, 24*, 73–81.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900–1902.

Hughes, K., & Coplan, R. J. (2010). Exploring processes linking shyness and academic achievement in childhood. *School Psychology Quarterly, 25*, 213–222.

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*, 88–110.

Kellam, S. G., Rebok, G. W., Ialongo, N., & Mayer, L. S. (1994). The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based preventive trial. *Journal of Child Psychology and Psychiatry, 35*, 259–281.

Ladd, G. W., & Profilet, S. M. (1996). The child behavior scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology, 32*, 1008–1024.

Maldonado-Carreno, C., & Votruba-Drzal, E. (2011). Teacher-child relationships and the development of academic and behavioral skills during elementary school: A within- and between-child analysis. *Child Development, 82*, 601–616.

Maziade, M., Côté, R., Boutin, P., Boudreault, M., & Thivierge, J. (1986). The effect of temperament on longitudinal academic achievement in primary school. *Journal of the American Academy of Child Psychiatry, 25*, 692–696.

McClowry, S. G. (1995). The development of the School-Age Temperament Inventory. *Merrill-Palmer Quarterly, 41*, 271–285.

McClowry, S. G. (2002). The temperament profiles of school-age children. *Journal of Pediatric Nursing, 17*, 3–10.

McClowry, S. G. (2014). *Temperament-based elementary classroom management.* Lanham, MD: Rowman & Littlefield.

McClowry, S. G., & Collins, A. (2012). Temperament-based intervention: Reconceptualized from a response to intervention framework. In R. Shiner & M. Zentner (Eds.), *Handbook of childhood temperament* (pp. 607–627). New York: Guilford Press.

McClowry, S. G., Rodriguez, E. T., Tamis-LeMonda, C. S., Spellmann, M. E., Carlson, A., & Snow, D. L. (2013). Teacher/student interactions and classroom behavior: The role of student temperament and gender. *Journal of Research in Childhood Education, 27*, 283–301. doi:10.1080/02568543.2013.796330

McClowry, S. G., Snow, D. L., & Tamis-LeMonda, C. S. (2005). An evaluation of the effects of *INSIGHTS* on the behavior of inner city primary school children. *Journal of Primary Prevention, 26*, 567–584.

McCroskey, J. C., & Daly, J. A. (1976). Teachers' expectations of the communication apprehensive child in the elementary school. *Human Communication Research, 3*, 67–72.

Nelson, L. J., Rubin, K. H., & Fox, N. A. (2005). Social and nonsocial behaviors and peer acceptance: A longitudinal model of the development of self-perceptions in children ages 4 to 7 years. *Early Education and Development, 20*, 185–200.

O'Connor, E. E., Rodriguez, E. T., Cappella, E., Morris, J. G., Collins, A., & McClowry, S. G. (2012). Child disruptive behavior and parenting sense of competence: A comparison of the effects of two models of *INSIGHTS. Journal of Community Psychology, 40*, 555–572.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48*, 85–112.

Rapport, M. D., Denney, C. B., Chung, K. M., & Hustance, K. (2001). Internalizing behavior problems and scholastic achievement in students: Cognitive and behavioral pathways as mediators of outcome. *Journal of Clinical Child Psychology, 30*, 536–551.

Raudenbush, S. W. (2009). Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings. *Journal of Education, Finance and Policy, 4*, 468–491.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T. (2011). *HLM 7, hierarchical linear and nonlinear modeling*. Chicago: Scientific Software International.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP″s impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Development, 82(1)*, 362–378.

Richmond, V. P., Beatty, M. J., & Dyba, P. (1985). Shyness and popularity: Children's views. *Western Journal of Speech Communication, 49*, 116–125.

Rimm-Kaufman, S. E., & Kagan, J. (2005). Infant predictors of kindergarten behavior: The contribution of inhibited and uninhibited temperament types. *Behavioral Disorders, 30*, 329–345.

Rudasill, K. M., & Konold, T. R. (2008). Contributions of children's temperament to teachers' judgements of social competence from kindergarten through second grade. *Early Education & Development, 19*, 643–666.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* London: Chapman and Hall.

Shapiro, E. S. (2004). *Workbook to accompany academic skills problems: Direct assessment and intervention* (2nd ed.). New York: Guilford.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69*, 493–525.

Slomkowski, C. L., Nelson, K., Dunn, J., & Plomin, R. (1992). Temperament and language: Relations from toddlerhood to middle childhood. *Developmental Psychology, 28*, 1090–1095.

Spere, K., & Evans, M. A. (2009). Shyness as a continuous dimension and emergent literacy in young children: Is there a relation? *Infant and Child Development, 18*, 216–237.

Webster-Stratton, C., Jamila Reid, M., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: evaluation of the incredible years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry, 49(5)*, 471–488.

Wenglinsky, H. (2004). Facts of critical thinking skills; what NAEP results say. *Educational Leadership, 62*, 32–35.

Zhang, Z., Zyphur, M. J., & Preacher, K. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods, 12*, 695–719.

Date Received: September 13, 2012
Date Accepted: August 9, 2013
Associate Editor: Shannon Suldo ■

Erin E. O'Connor, EdD, is an associate professor in the Department of Teaching and Learning at New York University. Dr. O'Connor's research examines the effects of teacher–child relationships on children's social–emotional and academic development, as well as the role of school-based interventions in improving relationships and student outcomes. She is one of the investigators of the Institute of Education Sciences study examining the efficacy of INSIGHTS for enhancing the academic learning context.

Elise Cappella, PhD, is an associate professor in the Department of Applied Psychology at New York University. Her scholarship focuses on understanding and enhancing classroom and school contexts as a means toward academic achievement and psychosocial development among students in urban schools. She directs the psychology and social intervention program at the New York University Steinhardt School and is one of the investigators of the Institute of Education Sciences study examining the efficacy of INSIGHTS for enhancing the academic learning context.

Meghan P. McCormick is an advanced doctoral student in the Department of Applied Psychology at New York University. Ms. McCormick's research examines the role of interpersonal relationships and social norms regarding children's development in school settings. She also studies how social–emotional learning can improve children's academic achievement in high-poverty schools.

Sandee G. McClowry, PhD, is a professor in the Department of Applied Psychology at New York University. She is the developer of the INSIGHTS into Children's Temperament intervention and was the principal investigator of the three clinical trials that tested its efficacy. Most recently, she was a Fulbright Scholar at the University of the West Indies.