Contemporary School Psychology
https://doi.org/10.1007/s40688-019-00270-z

Can Computers Teach Social Skills to Children? Examining the Efficacy of "The Social Express" in an African-American Sample

S. Kathleen Krach[1] · Michael P. McCreery[2] · Kanessa Miller Doss[3] · Dasha M. Highsmith[1]

Abstract
This study examined the efficacy of a computer-based social skills training program, *The Social Express*. Independent researchers evaluated the program at both a school-wide level (Tier 1) and at a referred group level (Tier 2). The sample included third-, fourth-, and fifth-grade students in a Title 1 public school with a 100% African-American population. At the Tier 1 level, pre-post (immediate) comparisons on a social skills rating scale indicated statistically significant differences by group at the $\alpha = .10$ level ($p = 0.058$). A significant Tier 1 quadratic effect for time (pre-test, post-test (immediate), post-test [delayed]) was found ($p = 0.029$) as well. At the Tier 2 level, pre-post comparisons indicated no statistically significant group improvement. Pre-post comparisons at the individual level found that about 39% of the children had statistically significant improvement in social skills, with 9% indicating a decrease in problem behaviors.

Keywords Technology · Social-emotional learning · Social skills training · School-based interventions · Positive behavioral interventions and supports (PBIS)

Children identified as needing support in the area of social skills are at higher risk of becoming adults with problems fitting into society (Babinski et al. 1999). Therefore, early social skills intervention is desirable, and schools have proven to be a good place to include social skills training programs (Alismail and McGuire 2015). Discussion of the research on social skills interventions should be examined through the lens of a comprehensive, integrated, three-tier model of prevention (Lane et al. 2009). For example, at Tier 1 (universal recipients), a meta-analysis of curriculum-based social skills training programs indicates an impressive level of success (Durlak et al. 2011). At a Tier 2 level (at-risk recipients), a meta-analysis specific to programs targeted at referred children with emotional and behavioral found less impressive gains (Maag 2006). Studies related to social skills training programs should be conducted at each of the appropriately tiered levels.

✉ S. Kathleen Krach
skrach@fsu.edu

[1] Department of Educational Psychology and Learning Systems, Florida State University, 1114 West Call Street, Tallahassee, FL 32306, USA

[2] Department of Teaching and Learning, University of Nevada Las Vegas, Las Vegas, NV, USA

[3] Department of Psychology, Troy University, Montgomery, AL, USA

Although there have been several studies on in-person social skills training programs, there have been fewer on the use of computer-based programs (Krach and McCreery 2016). The exceptions are studies specific to computer-based social skills training with children in special education (often seen as Tier 3) for a diagnosis of autism spectrum disorder (ASD). In a meta-analysis of these Tier 3, computer-based programs, Ramdoss and colleagues described most outcomes to be "unacceptable" (2012, p. 119).

It is unknown if these same computer-based, social skills training programs would result in behavioral changes at either a Tier 1 or Tier 2 level. It is important to know if computer-based social skills programs would work at these other levels, as this may ease the instructional burdens of teachers trying to differentiate instruction (Cobb 2010). Additional research must determine the efficacy of computer-based social skills training programs for both a school-wide (Tier 1) and a referred (Tier 2) population of children.

Social Skills Training

Positive Behavioral Interventions and Supports

In the past two decades, schools have begun using a Multi-Tiered System of Supports (MTSS; Benner et al. 2013) or the

Published online: 13 January 2020

more structured Comprehensive, Integrated, Three-Tier model of prevention (CI3T, Lane et al. 2009, 2014). Both of these models include supports for children with social, emotional, and behavioral needs through a framework called Positive Behavioral Interventions and Supports (PBIS; PBIS National Technical Assistance Center 2009). Embedded in the PBIS framework is the need for research-based interventions. Tier 1 tends to focus on improving ways in which the school system deals with broad-based social, emotional, or behavioral issues (Bradshaw et al. 2008; Luiselli et al. 2005; March and Horner 2002). At Tier 2 and Tier 3, the interventions tend to focus on small groups (Chafouleas et al. 2007; March and Horner 2002) or individualized help (Brooks et al. 2003; Scott et al. 2005) respectively.

Social Skills Training and Social Thinking

Historically, PBIS often concentrates on encouraging and increasing positive behaviors instead of just decreasing negative behaviors (Sugai and Simonsen 2012). Specifically, the literature on PBIS addresses the development of positive social skills or pro-social behaviors across each intervention tier (Sugai 2007; Sugai and Horner 2006; Sugai and Simonsen 2012). Sansosti (2010) specifically described social skills programs that fit well across each of the three tiers of the PBIS framework. According to their work, Tier 1 social skills programs often include character education and social skills training. Tier 2 programs include social skills groups and peer-mediated approaches, and Tier 3 programs consist of programs such as social stories and power cards.

Winner's *Social Thinking* One popular social skills curriculum is Winner's (2005) *Social Thinking* cognitive-behavioral program. Winner (2005) designed *Social Thinking* to be used by professionals and parents to teach social skills and problem solving. Winner (2005) describes her model with the acronym ILAUGH for (I)mitation of Language, (L)istening with Eyes and Brain, (A)bstract and Inferential Language/Communication, (U)nderstanding Perspective, (G)estalt Processing/Getting the Big Picture, and (H)umor and Human Relatedness. An article by Winner and Crooke (2009) provides research citations supporting the separate components of the program. Additional empirical support exists for the *Social Thinking* program as a whole (Crooke et al. 2008; Koning et al. 2013; Lee et al. 2015).

Studies on the *Social Thinking* curriculum indicate its effectiveness within a tiered PBIS framework at the Tier 2 level. Crooke et al. (2008) found significant pre-post treatment improvements in both verbal and non-verbal social behaviors exhibited by children after experiencing an 8-week, small-group version of the program. Lee et al. (2015) found statistically significant improvement pre-post intervention across ratings of abstract and concrete language, empathy, and non-verbal communication after 12 weeks of small-group instruction. Finally, Koning et al. (2013) found statistically significant improvements in children's interactions with peers as well as perceptions and knowledge of social situations following a 15-week, group-based program. At the Tier 3 level, the *Social Thinking* program has been found effective when used with six students diagnosed with ASD (Crooke et al. 2008) and four other students with the same diagnosis in a different study (Lee et al. 2009). No studies could be identified for use of *Social Thinking* at the Tier 1 level.

The *Social Thinking* program is not without controversy. Leaf et al. (2016) lay forth a solid argument that *Social Thinking* is a "pseudoscience" program and should not be seen as best practice, but the authors of the *Social Thinking* program dispute aspects of this argument in their position paper (Crooke and Winner 2015). The purpose of the current paper is not to determine the effectiveness of the *Social Thinking* program; instead, it is to evaluate the effectiveness of a computer-based program that was based off of the ideas embedded within the *Social Thinking* program.

Computer-Based Social Skills Training

The Social Express Developed by The Language Express (2013), *The Social Express* is a proprietary social skills training software based on Winner's *Social Thinking* program (2005). According to *The Social Express* website (http://thesocialexpress.com), the purpose of the software is to provide school-aged children with the opportunity to practice real-life social interactions and learn foundational socio-emotional skills. Due to the proprietary nature of the software, the study authors (who are unconnected to the makers of the program) were able to report on neither the technical designs nor the specifics of how the intervention was evaluated during the design process.

According to *The Social Express* teaching manual (The Language Express 2013), the program uses vignettes and curriculum materials to teach behavioral learning principles in terms of "what," "how," and "why." Thus, learning is broken down into the comprehension of concepts (what), how those concepts are expressed in social situations, and why those pro-social expressions are important. The publishers of *The Social Express* (The Language Express 2013) state that this is accomplished through the scaffolded instruction of targeted skills sets, including language (i.e., listening and rule following), logic (i.e., critical thinking and attention), and regulation (i.e., emotional and self-management).

In addition to many private and public treatment providers, several school districts have adopted *The Social Express.* These include the Los Angeles Unified School District, Hartford School District, Lennox School District, and McKinley School District. The software program was a 2013 winner of the District Administration Magazine's "Reader's

Choice Top 100 Products" (District Administration 2013). According to the software's website, *The Social Express* has won positive recognition from several software and application reviewers across the internet. Although it has been well received, at the time of this study, there are no independent, published efficacy studies to support the use of the program.

The only research available on the efficacy of this program is a dissertation by Kelly (2015) that examined the use of *The Social Express* with a group of children who had diagnoses of high-functioning autism spectrum disorder or Asperger's (Tier 3). The study questioned whether adding a computer-based program (*Th Social Express*) to a traditional social skills training program resulted in children's overall lower scores on an autism rating scale (*Asperger Syndrome Diagnostic Scale, ASDS*; Myles et al. 2001) and higher scores on a social skills rating scale (*Social Responsiveness Scale, Second Edition*; *SRS-2;* Constantino and Gruber 2005) when compared to children who only received the traditional training program. Findings indicated that the group whose training program included *The Social Express* showed a non-significant decrease on the autism rating scale when compared to the traditional group. For social skills, subjects did not show a significant increase on the SRS-2 subscales of Social Awareness, Social Cognition, Social Communication, or Autistic Mannerisms; however, there was a statistically significant increase in the Social Motivation subscale.

With the exception of this one dissertation study, the makers of *The Social Express* have depended on the research on the *Social Thinking* curriculum to support the overall use of the program. However, there are several potential problems with generalizing previous research supporting *Social Thinking* to *The Social Express*. For example, there are no studies extant on the use of the *Social Thinking* curriculum as a Tier 1 intervention. In addition, there are no studies on a referred population other than those diagnosed with autism spectrum disorder. Finally, there are no studies on the use of *Social Thinking* tasks in an alternative format (i.e., computer-based). Thus, additional research is needed to examine the efficacy of *The Social Express* beyond its association with the *Social Thinking* framework. The current study sought to examine the effects of a computer-based program (*The Social Express*) at both Tier 1 and Tier 2 levels of a PBIS, multi-tiered framework.

Research Questions

Tier 1 Study Do social skill rating scale data indicate a significant pre-post test score difference following the completion of *The Social Express* for an entire school population (Tier 1) when comparing control and experimental groups?

Tier 2 Study Do social skill rating scale data indicate a significant pre-post test score difference following the completion of *The Social Express* for a referred population of children with behavior problems (Tier 2)?

General Procedures

Intervention Materials (Tier 1 and Tier 2)

The Social Express (The Language Express, 2013) is a vignette-based computer program used to teach social skills to children at their own pace. All vignettes provide contextualized visual and auditory stimuli that model social scenarios. Children watch animated vignettes on a computer and are then asked to select the correct pro-social behavioral choice associated with that scenario. Based on their choice, they then watch an animated clip that shows the result of that choice. If a non-pro-social choice is made, the software presents the correct pro-social choice with a description of why it was correct. Then the child is shown an animated clip demonstrating the correct choice and its impact on the social situation. At the end of each vignette or set of vignettes, the child is then given a multiple-choice quiz to determine how much was learned about the concepts discussed.

The current version of *The Social Express* consists of 81 animated interactive scenarios designed to model socio-emotional skills in a variety of areas. Included in these are conflict resolution, critical thinking, relationship skills, and self-management. However, at the time of data collection, *The Social Express* only had 24 vignettes available. In addition to these computer-based program materials (e.g., quizzes and vignettes), there were teacher-provided curriculum materials to accompany the webinars (if requested).

Procedures (Tier 1 and Tier 2) This project comprised two studies in the course of 2 years. The Tier 1 study took place during year one, and the Tier 2 study took place in the second year. The first year's group consisted of the entire K-5 school (Tier 1 study); however, only the third-, fourth-, and fifth-grade teachers were compliant with the program. The second year's group (Tier 2 study) consisted only of referred children in the third, fourth, and fifth grades. Given the anonymous nature of each study, and the fact that the study took place over two consecutive years, there is a small chance that some students (not identified) participated in each study. None of the participants had access to both Tier 1 and Tier 2 services at the same time.

Participants (Tier 1 and Tier 2) Participants for both studies were from the same Title 1 elementary school in a mid-sized city in the Southeastern United States. This school was chosen because it serves as a professional development school for the local university. No other research projects were going on concurrently at the school. The school participants (teachers

and students) consisted 100% of non-Hispanic, African-American students. No other races or ethnicities were represented; this reflects the population of the school. Each classroom had two to four computers available for student use, including internet connection and headphones. Study and consent procedures were approved by IRB.

Tier 1—Study 1

Methods

Procedures In the Tier 1 study, all children were provided the opportunity to use the social skills training program, *The Social Express*; however, only students whose parents provided consent were included in the study. The research team trained the teachers how to use the program in a 30-min workshop during a teacher preparation day early in the fall of the school year.

After training was complete, researchers randomly divided classes into two groups (experimental/fall and control/spring) by picking teacher's names at random from slips of paper sorted by grade. If there were an even number of classes in a single grade, then half of the classes went into each group. If there was an odd number of classes in a single grade, then the extra class went into the experimental group.

Both groups received the program, one in the fall and one in the spring The fall/experimental group did not have access to the program during the spring (access was only available in the fall). The spring/control group did not have access to the program in the fall (access was only available in the spring). Teachers were not informed which group (fall or spring) constituted the control group and which was the experimental group.

Surveys were collected three times during the entire year. Pre-test, teacher rating scale data were collected for both groups before program access was granted for the experimental group in the fall (around the first of October) and immediately after the program access was withdrawn (post-test (immediate)) around the middle of December. Rating scale data was collected for both groups at the end of the entire Tier 1 project (post-test (extended)) around the first week of May.

Tier 1 lessons were open and available for 10 weeks. The use of a 10-week program was selected based on recommendations from the US Department of Education; these recommendations stated that structured learning of new material should be reviewed for a period between several weeks and several months (Pashler et al. 2007). The first 6 weeks consisted of three vignette lessons each week. Weeks seven and eight consisted of four vignette lessons. Weeks nine and ten were assigned for catching up on any sessions not completed.

Teachers were asked to direct the children to do the assigned lessons when they had free time to do so. Children completed the vignettes independently (with headphones) and took online quizzes at the end of each vignette. Vignettes were interactive; children watched a short video, were offered a choice, and received feedback on that choice. In addition, at the end of each lesson, children completed a quiz to determine what they learned. No psychometric data is available on the quiz. No data from the child's vignette interactions were made available to the researchers by the software manufacturers. Only quiz total results were provided by the publisher; therefore, no psychometric assessment of these quiz results could be conducted for the current study.

A school psychologist came every week in the fall to provide the classes with support on both the technology and the program. In order to measure treatment integrity, all teachers were interviewed about the study each week; in addition, a paper survey was provided at the end of the program. None of the additional social skills curriculum materials were used by either the teacher or the school psychologist. The only social skills instruction for the Tier 1 study came from the software program. Although the spring group had access to the program they did not receive the additional weekly reminders or technical support that the fall group received.

Participants It is not uncommon in school-based research for samples to decrease in number due to a lack of parental consent, as well as teacher non-compliance in use of the intervention and/or providing data (Owens and Murphy 2004). These issues held true for the current study. Out of the 517 total students in the school, the parents of 428 students affirmed consent to use their child's data for research purposes: experimental (original $n = 260$) and control (original $n = 168$).

Before analysis, experimental group data were removed for the following reasons: insufficient number of sessions (10 or fewer vignettes completed; $n = 76$), no sessions completed ($n = 136$ participants), and lack of pre and/or post data ($n = 7$). Following the removal of these participants, a review of the remaining students found no kindergarten or second grade participants remaining and only one student in first grade. This single student was removed from the sample. This left a final experimental sample of 38 participants. Before analysis, control group data were removed for the following reasons: lack of pre and/or post data ($n = 23$) and lack of matching experimental data in those grades (i.e., Kindergarten, first, or second grade; $n = 78$). Finally, all students in the control group who did not participate in the program in the Spring were removed prior to the final analysis. This left a sample of 34 control group participants.

An independent sample t test was conducted to compare pre-post (immediate) difference scores between the students who were included with those who were excluded from the sample (when data were available and consent provided).

Results indicated no significant difference for those included in the sample when compared to those excluded from the sample, $t(361) = 1.178$, $p = 0.240$, $d = 0.004$.

The Tier 1 total sample ($n = 74$) consisted of 47 third-grade, 10 fourth-grade, and 17 fifth-grade students. The final total gender breakdown included 33 females and 41 males. The racial makeup of the final sample consisted of 74 African-American students. A power analysis using G*Power 3.0.10 (Erdfelder et al. 1996) set for an effect size = 0.25, $\alpha = 0.05$, and power (1-β) = 0.95 resulted in the need for 44 individuals in the sample. Therefore, the sample size should be viewed as sufficient.

Instruments The same instrument was used for each of the time markers in the study (i.e., pre-intervention, immediately post intervention, and 5-month post-intervention). Separate measures were used for fidelity assessments.

Study Instruments For the Tier 1 portion of the study, the instrument needed to provide a quick method of collecting data. This is because data were collected on all students in each class multiple times by every teacher. Tier 1 data were collected from each child's homeroom teacher on the Pro-Social Behavior subscale from the Performance Screening version of the Social Skills Improvement System (SSIS-PSG; Gresham and Elliot 2008). The SSIS-PSG was selected because it is a scale commonly employed by teachers due to ease of use (Kelly 2015). The Pro-Social Behavior subscale of the SSIS-PSG consists of an item that asked teachers to rate students on a 1–5 scale, where "1" states that the student possesses very limited pro-social skills and "5" indicates that he/she possesses excellent pro-social skills. The manual (Gresham and Elliot 2008) reports reliability data in the form of test-retest reliability (.60–.70) and interrater reliability (.55–.68) to meet acceptable levels for teacher-based assessment (Landis and Koch 1977). Independent assessment of the scale found solid psychometric qualities for both reliability and validity (Krach et al. 2016).

Fidelity Instruments Treatment fidelity was measured through weekly teacher interviews where the teachers were asked the following question: Is your class using the program? An additional paper survey was given to the teachers at the end of the study. The survey included the following question regarding fidelity: did you do the lessons on the dates suggested by [name of researcher]? Although the responses were open-ended, they all fell into one of five categories: yes, mostly yes, some, mostly no, and no. Finally, fidelity was also measured through the examination of backend data from *The Social Express* from the publisher of the software program.

Data Analyses An independent sample *t* test was run on the pre-test data for both final groups (experimental and control) to ensure that no pre-existing differences between the groups were found. In addition, an assessment of treatment fidelity was run to ensure that the individuals in the experimental group received the intervention. Finally, a mixed ANOVA was used to determine within-subjects differences in pre-test/post-test (immediate and extended) for social skills ratings between experimental and control groups.

Results The first research question asked if rating scale data indicated a significant pre-test/post-test score difference following the completion of *The Social Express* for an entire school population (Tier 1). Homeroom teachers provided pre-test/post-test SSIS: Pro-Social Behavior (SSIS - PSG: PSB; Gresham and Elliot 2008) data for all consented students. Table 1 provides descriptive statistics for these scores.

An independent sample *t* test was run to examine control/experimental group differences in pre-test, SSIS-PSG: PSB scores prior to the intervention. The Levene's test for equal variance indicated that equal variances can be assumed ($F = 0.028$, $p = 0.867$). Results indicated no significant pre-test score differences, $t(72) = 0.597$, $p = 0.552$, between experimental and control groups on the SSIS: PSB.

Tier 1 treatment fidelity were as follows. During the weekly fidelity interview checks, 100% of the teachers reported that their classes were using the program; this was discrepant from the backend data derived from the software where only 32.56% of the students in the classes were using the program. The paper survey was also discrepant from the backend data derived from the software. Although eight teachers stated that they used the program with fidelity, the average percent of their students who actually used the program was 59.02%. Three of the 12 teachers who said that they used or mostly used the program did not have a single student use the program. Therefore, it can be assumed that the fidelity of the program is best evaluated by the backend data from the software. Only students who had backend data that verified their actual involvement in the program were included in this study.

Table 1 Tier 1 study: Descriptive statistics

Pro-social behavior	Group	*M*	SD	*N*
Pre-test	Experimental	3.71	1.228	38
	Control	3.56	1.186	34
	Total	3.64	1.202	72
Post-test (immediate)	Experimental	4.11	.953	38
	Control	3.71	1.315	34
	Total	3.92	1.148	72
Post-test (extended	Experimental	3.63	.970	38
	Control	3.65	1.412	34
	Total	3.64	1.190	72

The Pro-Social Behavior item on the SSIS-PSG is on a scale of 1 (very limited/extreme difficulty/poor control) to 5 (excellent skills)

A mixed-design ANOVA with a within-subjects factor of time (pre-intervention, immediately post-intervention, and 5-month post-intervention) and a between-subjects factor (experimental or control) was used to analyze the data in this study. Mauchly's test indicated that the assumption of sphericity had not been violated ($\chi^2(2) = 0.488$, $p = 0.784$). Box's test of equality was significant, M (6.34295) = 15.028, F = 2.387, $p = 0.026$, indicating that the means for the groups were not equal. Levene's Test of Equality found that pre-test data (F (1, 70) = 0.214, $p = 0.645$) and post-test (immediate) (F (1, 70) = 3.064, $p = 0.084$) data met the assumptions of equality; however, post-test (extended) data did not (F (1, 70) = 7.911, $p = 0.006$) Therefore, ANOVA findings derived from the post-test (extended) data may need to be considered with caution.

Main linear effects of time, F (1, 70) = 0.001, $p = 0.974$, partial $\eta^2 < 0.001$ and time*group, F(1, 70) = 0.346, p = 0.558, partial $\eta^2 = 0.005$, were not significant.

Main quadratic effects of time, F (1, 70) = 4.966, $p = 0.029$, partial $\eta^2 < 0.066$, was significant; main quadratic effects of time*group, F(1, 70) = 1.889, $p = 174$, partial $\eta^2 = 0.026$, were not significant. Figure 1 provides a visual for these findings.

Given the significant quadratic effects, a second linear analysis was run for pre-test and post-test (immediate) to evaluate if the initial growth found was statistically significant. The main linear effects, F (1, 73) = 3.702, $p = 0.058$, partial $\eta^2 = 0.048$, were not significant at the 0.05 level, but were significant at the 0.10 level.

Discussion At the Tier 2 level, pre-post evaluation found no statistical significance for the whole group; the effect size was small. At the individual participant level, there is always an expectation that about 33% will fail to respond to Tier 2 treatment and will need to go on to Tier 3 services; therefore, any program wherein 66% of participants demonstrate improvement should be considered efficacious (Sugai and Horner 2006). This 66% threshold was not met in the findings from the current study. For the area of social skills, about 40% showed significant improvement; only 9% showed significant improvement in the area of problem behaviors.

A post-analysis examined those who did better, worse, or about the same in their response to the intervention. Those who demonstrated that they were learning the content by performing well on the in-program quizzes were also the ones more likely to fall in the "some improvement" or "statistically significant improvement" groups. This indicates that *The Social Express* may help some children at a Tier 2 level but not others. What determines who does, and does not, benefit may be related to an, as yet unknown, intra-child variable (e.g., attention to the task, intellectual ability). Further analysis is warranted at the Tier 2 level examining the program along with any intra-child variables.

Tier 2—Study 2

Methods

Procedures The second study (Tier 2) consisted of more expanded instruction in a small-group setting. Referred students were placed in groups of five to eight students; each grade level had two different groups. Each session was led by an educator, a school mental health professional, and a graduate assistant. Both the computer-based vignettes and the social skills in-person curriculum materials provided by *The Social Express* were used in the Tier 2 program. The in-person curriculum was designed by the program publisher and is proprietary. As with the computer-based curriculum, the paper-based materials are also based on the *Social Thinking* curriculum (Winner 2005).

Each session began with an introductory activity to activate prior knowledge by reviewing previous vignettes and explore

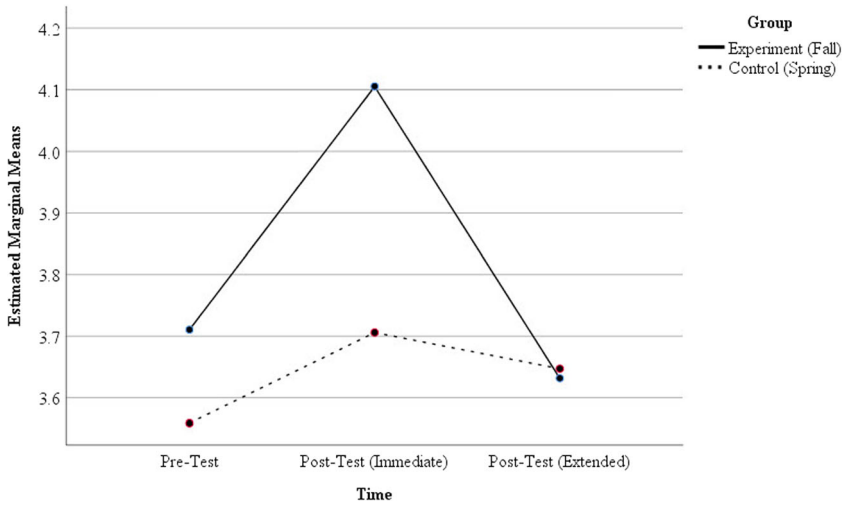

Fig. 1 Visualization of means across data collection times and experimental groups

possible connections with the current vignette's title. The vignette consisted of interactive lessons that allowed the students to select the actions of the avatar and watch the consequences. During each 40-min session, the entire group completed one vignette for approximately 5 min per session together as a unit; there were two sessions per week across 10 weeks. A program length of 10 weeks was selected based on recommendations from the RTI Action Network (RTI-AN; Metcalf 2015); the RTI-AN recommend a minimum of 6 weeks with an average of eight to 12 weeks for Tier 2 programs. In addition to the vignettes, the publisher of the software program provided supplemental instructional materials with activities to reinforce key concepts (e.g., designing comic strips, creating posters, storyboarding). At the end of each session, children individually completed the online quiz related to that week's topic area.

Participants For the Tier 2 portion of the study, teachers were asked to refer the top 10% of children in their class exhibiting poor social skills who were not currently receiving special education services in the Other Health Impaired (OHI) category for Attention Deficit Hyperactivity Disorder (ADHD), the autism category, or the emotionally disturbed (ED) category under IDEA (2004). These students were not included because this additional time being removed from their classrooms could violate their least restrictive environment requirement (IDEA, 2004). The average number of office referrals per child was 3.33 (SD = .4.488) for that school year. All of the subjects were non-Hispanic and African American. All Tier 2 group participants (n = 42) had parental consent to be included in the study.

The Tier 2 students who were removed from the data-set prior to analysis had the following issues: no post-test data were available (n = 2), the validity (F) index score on the rating scale (SSIS-SRS; Gresham and Elliot 2008) indicated "caution" on pre-test data (n = 2) or post-test data (n = 2), extreme outlier scores (n = 1), and/or non-compliance in completing at least 10 sessions (n = 12). This resulted in a total usable sample of 23 subjects; 10 of these were female (43.5%) and 13 were male (56.5%). There was an even distribution of nine children in each grade. Each grade consisted of two groups; the average number of sessions was 12.78 (SD = 2.730). All of the students were African American.

Instruments Because the program was provided directly by the researchers, no fidelity checks were needed.

Study Instruments With fewer children included in the Tier 2 sample, a lengthier data collection tool could be used. Tier 2 data were collected from teacher ratings on the Social Skills Rating System version of the Social Skills Improvement System (SSIS-SRS; Gresham and Elliot 2008). The SSIS-SRS has a Social Skills composite score consisting of subscales in communication, cooperation, assertion, responsibility, empathy, engagement, and self-control. This Social Skills composite has demonstrated internal consistency data from the manual; coefficient alpha was .97 (Gresham and Elliot 2008). In addition, there is a Problem Behaviors composite comprising externalizing, bullying, hyperactivity/inattention, and internalizing subscales. The Problem Behaviors composite has demonstrated internal consistency data from the manual; coefficient alphas for ages 5–12 of .95 (Gresham and Elliot 2008). Validity scales determined score usability based on results from the F-Index (Faking Bad).

Data Analyses Because no control group was available for the Tier 2 portion of the study, alternate methods were used to determine the response to intervention. First, researchers calculated a comparison value to determine statistically significant differences using Anastasi and Urbina's (1997, p. 111) formula (see Fig. 2). This value was then compared to individual pre-post difference scores to determine if statistically significant growth occurred for individual students. In addition, a paired sample t test with effect size change statistics were calculated for the entire group.

Results The second research question asked if rating scale data indicated a significant pre-test/post-test score difference following the completion of *The Social Express* for a referred population of children with behavior problems (Tier 2). Homeroom teachers provided SSIS-SRS (Gresham and Elliot 2008) data for all consented students. Table 2 provides descriptive statistics for these scores.

Frequency data on pre-test/post-test change on the Social Skills Composite scores and Behavioral Problems Composite scores can be found in Table 3. To determine the statistical significance comparison value using Anastasi and Urbina's (1997) formula, the reliability coefficients described above were taken directly from the manual (Gresham and Elliot 2008). Raw score standard deviations were calculated from the pre-test and post-test data (SD = 1.585 for Social Skills composite and SD = 4.9 for the Behavioral Problems composite). The comparison values for the Social Skills composite was 0.76 for the raw scores. The comparison values for the Behavioral Problems composites was 3.04 for the raw scores.

Table 3 provides data on the frequency of subjects whose differences scores met or exceeded this comparison value, and thus experienced statistically significant change. From the comparison score data, it is clear that pre-post test difference scores indicated worse behaviors for 39.1% of the referred student for social skills and 34.8% for problem behaviors. Changes in raw scores indicated that 39.1% of the subjects showed statistically significant increases in social skills and 8.7% showed statistically significant differences in decreases in problem behaviors. Follow-up analysis using a Between-Subjects ANOVA found that raw scores on the within-

$$SE_{diff.} = SD\sqrt{2 - r_1 - r_2}$$

- SD = Standard deviation of the score
- r_1 = reliability of 1st test
- r_2 = reliability of 2nd test

SE_{diff} X 1.96 provides probability of chance

Fig. 2 Statistically significant differences formula Anastasi and Urbina (1997, p. 111)

program quizzes (measuring understanding of the content) were statistically significantly different at the 0.05 level across group performance labels (i.e., worse, same, some improvement, statistically significant improvement) for both Social Skills ($F(2, 14) = 4.115$, $p = 0.039$, partial $\eta^2 = 0.370$) and Problem Behaviors ($F(23, 14) = 3.983$, $p = 0.30$, partial $\eta^2 = 0.460$).

Tier 2 data were also examined for the whole group using a paired samples *t* test for statistical significance and effect size. The statistical significance calculation should be noted as problematic because there is not a large enough sample to have enough power; however, the effect size may still be considered. For the Social Skills composite score, pre-test/post-test difference findings for raw scores was $t(22) = -0.535$, $p = .0.598$, $d = 0.116$. For the Problem Behaviors composite, pre-test/post-test difference findings for raw scores was $t(22) = .381$, $p = .707$, $d = 0.109$.

Discussion At the Tier 1 level, findings indicate that students receiving the intervention demonstrated improved social skills at a level approaching significance ($p = .058$) when compared to the control group. Additional findings indicate that this improvement does not maintain. Specifically, quadratic analysis and visual inspection of the data shows that, although children receiving the intervention experienced some improvement immediately following completion of the intervention, any improvements were lost within 5 months once the intervention was removed.

Tier 1 and Tier 2 Discussion

Positive Behavioral Interventions and Supports (PBIS) strategies can be used to help students develop social skills (Sansosti 2010; Sugai 2007; Sugai and Horner 2006; Sugai and Simonsen 2012). Although social skills training programs are traditionally taught in-person, there are computer-based programs available to teach socio-emotional learning skills. The purpose of the current research was to evaluate one such programs, *The Social Express*, at both the Tier 1 and Tier 2 levels within a PBIS framework. *The Social Express* was chosen due to its popularity (District Administration 2013) and the fact that it lacks independent evaluation data; the researchers for the current study received no compensation.

This study should be considered exploratory in nature. This is because there were issues with the sample; these issues were not unexpected given prior research on using school-based samples. The two specific issues, sample attrition and teacher non-compliance in data collection, are described as common for school-based samples by Owens and Murphy (2004). Both of these issues were present in this study. In addition, there were generalization issues because the current study only examined an upper-elementary and Title-1 school population who were exclusively African-American. There was a potential issue regarding the selected instrument used in the Tier 1 study. The shorter instrument was chosen because teachers were asked to complete the ratings multiple times for every child in every classroom. Although more items would provide a more psychometrically sound study, a longer instrument would have been impractical in this setting.

Finally, the time limitation of 10 weeks for both the Tier 1 and the Tier 2 studies may not determine true efficacy.

Table 2 Tier 2 study: Descriptive statistics

	Pre-test			Post-test		
	SSIS-SSRS: composite			SSIS-SSRS: composite		
	N	*M*	*SD*	*N*	*M*	*SD*
Social skills raw	23	10.479	2.108	23	10.739	2.359
Social skills SS[a]	23	40.261	0.541	23	40.478	0.846
Problem behaviors raw	23	9.783	1.380	23	9.6087	1.777
Problem behaviors SS[b]	23	88.217	1.565	23	88.044	1.988

SSIS-SSRS composite scores were provided as raw scores and standard scores (SS)

[a] Standard score (SS) for normative group for social skills: $M = 50$, $SD = 10$

[b] Standard score (SS) for normative group for problem behaviors: $M = 100$, $SD = 15$

Table 3 Tier 2 (referred group) improvement frequencies

	No improvement		Improvement	
Tier 2 referred group	Worse[b]	Same	Some[b]	Statistically significant[a]
Social Skills raw	9 (39.1%)	5	0	9 (39.1%)
Problem Behaviors raw	8 (34.8%)	7	6	2 (08.7%)

$n = 24$ across all scores

[a] The SSIS-SSRS composite difference scores were generated by subtracting pre-test scores from post-test scores. These scores were then compared to a value calculated by the formula found in work by Anastasi and Urbina (1997, p. 111, Fig. 2). Any difference score for an individual that was higher than the comparison value was considered here to be a statistically significant improvement. The comparison values for the Social Skills composite was 0.76 (rounded to 1.0) for the raw scores. The comparison values for the Behavioral Problems composites was 3.04 (rounded to 3.0) for the raw scores

[b] Any scores that showed a non-statistically significant improvement were rated here as "some." Any that went in a negative direction was rated here as "worse"

Although the restriction was based on recommendations from solid sources (Metcalf 2015; Pashler et al. 2007), those sources did not provide evidence of support for their recommendations. Interventions have traditionally been recommended for between 8 and 10 weeks, but this does not mean we know what the ideal time period for maximum effect might be. Additional research using *The Social Express* for a longer period of time is warranted. In addition, given that most of these study-related issues occurred because of research restrictions within a public school setting, any future study on this program may need to be done in a controlled laboratory environment.

For the most part, this study evaluated the efficacy of one particular software program. However, it may also be seen as an evaluation of other embedded concerns. For example, the makers of *The Social Express* describe the program as based off of the *Social Thinking* training program (Winner 2005). Therefore, in part, this study examined the effectiveness of that curriculum. The current study also examined the use of technology to teach social skills; this should be considered separately from the efficacy of the curriculum (Ramdoss et al. 2012). Finally, the software program's use of vignette-based training programs for social skills may constitute its own area of study (Shukla-Mehta et al. 2010). Unfortunately, the current study did not allow for a separate assessment of each of these concerns. That is in part due to the fact that *The Social Express* consists of so many different innovations.

The current findings using a computer-based, social skills training program fall in line with research on more traditional, in-person programs. As was found by Durlak et al. (2011) for traditional programs, this computer-based program demonstrated some level of success at the Tier 1 level. As was found by Maag (2006) for traditional programs, this computer-based program was much less successful for children at the Tier 2 level. The current study provided some insight into the efficacy of *The Social Express* (Sugai and Horner 2006). However, more study on program efficacy is needed before generalizing these findings beyond an African-American sample.

Funding This study was funded by a material grant from The Social Express. The grant consisted of the option to use (free of charge) the software program in one, Title-1 elementary school for 2 years. No funds were provided. No requirements were placed on the researchers other than the software be used as part of a research project.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Research Involving Human Participants and/or Animals IRB approval was granted for both Troy University and Florida State University for this project.

Informed Consent All children had access to the software programs. Parents provided written informed consent for researchers to be able to use their child's/children's data in this study.

References

Alismail, H. A., & McGuire, P. (2015). Twenty-first century standards and curriculum: current research and practice. *Journal of Education and Practice, 6*(6), 150–154 Retrieved from https://files.eric.ed.gov/fulltext/EJ1083656.pdf.

Anastasi, A., & Urbina, S. (1997). *Psychology testing*. Upper Riverside: Prentice Hall.

Babinski, L. M., Hartsough, C. S., & Lambert, N. M. (1999). Childhood conduct problems, hyperactivity-impulsivity, and inattention as predictors of adult criminal activity. *Journal of Child Psychology and Psychiatry, 40*(3), 347–355.

Benner, G. J., Kutash, K., Nelson, J. R., & Fisher, M. B. (2013). Closing the achievement gap of youth with emotional and behavioral disorders through multi-tiered systems of support. *Education and Treatment of Children, 36*(3), 15–29. https://doi.org/10.1353/etc.2013.0018.

Bradshaw, C. P., Koth, C. W., Bevans, K. B., Ialongo, N., & Leaf, P. J. (2008). The impact of school-wide positive behavioral interventions and supports on the organizational health of elementary schools. *School Psychology Quarterly, 23*(4), 462–473. https://doi.org/10.1037/a0012883.

Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase

academic engagement and work completion. *Journal of Positive Behavior Interventions, 8*, 144–152. https://doi.org/10.1177/10983007030050030301.

Chafouleas, S., Rile-Tillman, C., Sassu, K., LaFrance, M., & Patwa, S. (2007). Daily behavior report cards: an investigation of the consistency of on-task data across raters and methods. *Journal of Positive Behavior Interventions, 9*, 30–37. https://doi.org/10.1177/10983007070090010401.

Cobb, A. (2010). To differentiate or not to differentiate? Using internet-based technology in the classroom. *The Quarterly Review of Distance Education, 11*(1), 37–45.

Constantino, J. N., & Gruber, C. P. (2005). *Social responsiveness scale, second edition (SRS-2)*. Torrance: Western Psychological Services.

Crooke, P. J., Hendrix, R. E., & Rachman, J. Y. (2008). Brief report: measuring the effectiveness of teaching social thinking to children with Asperger syndrome (AS) and high functioning autism (HFA). *Journal of Autism and Developmental Disorders, 38*(3), 581–591. https://doi.org/10.1007/s10803-007-0466-1.

Crooke, P. J., & Winner, M. G. (2015). *Research to frameworks to practice: social thinking's layers of evidence*. Santa Clara: Thinking Social.

District Administration (2013). Readers' choice top 100 products for 2013. *District Administration Magazine*. http://www.districtadministration.com/top-100-winners-2013?keys&items_per_page=10&page=1

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: a general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*(1), 1–11. https://doi.org/10.3758/BF03203630.

Gresham, F., & Elliot, S. N. (2008). *Social skills improvement system (SSIS) rating scales*. San Antonio: Pearson Assessments.

Koning, C., Magill-Evans, J., Volden, J., & Dick, B. (2013). Efficacy of cognitive behavior therapy-based social skills intervention for school-aged boys with autism spectrum disorders. *Research in Autism Spectrum Disorder, 7*, 1282–1290. https://doi.org/10.1016/j.rasd.2011.07.011.

Kelly, R. M. (2015). Group social skills interventions for children with Asperger's: the effects of parent-guided social skills software. (doctoral dissertation). Retrieved from https://scholarcommons.sc.edu/etd/3722

Krach, S. K., & McCreery, M. P. (2016). Technology and positive behavioral support: Evaluation, selection, and implementation of computer-based socio-emotional training in schools. In S. Y. Tettegah & D. Espelage (Eds.), *Emotions and technology: Communication of feelings for, with, and through digital media – Volume I: Emotions, learning, and technology* (pp. 159–174). Waltham: Elsevier.

Krach, S. K., McCreery, M. P., Wang, Y., Mohammadiamin, H., & Cirks, C. K. (2016). Diagnostic utility of the social skills improvement system performance screening guide (SSIS-PSG). *Journal of Psychoeducational Assessment, 35*(4), 391–409. https://doi.org/10.1177/0734282916636500.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310.

Lane, K. L., Kalberg, J. R., & Menzies, H. M. (2009). *Developing schoolwide programs to prevent and manage problem behaviors: a step-by-step approach*. New York: Guilford Press.

Lane, K. L., Oakes, W. P., & Menzies, H. M. (2014). Comprehensive, integrated, three-tiered models of prevention: why does my school - and district - need an integrated approach to meet a students' academic, behavioral, and social needs? *Preventing School Failure: Alternative Education for Children and Youth, 58*(3), 121–128. https://doi.org/10.1080/1045988X.2014.893977.

Lee, K. Y., Crooke, P. J., Lui, A. L., Kan, P. P., Mark, Y. M., van Hasselt, C. A., & Tong, M. C. (2015). The outcome of a social cognitive training for mainstream adolescents with social communication deficits in a Chinese community. *International Journal of Disability, Development and Education, 63*(20), 1–23. https://doi.org/10.1080/1034912X.2015.1065960.

Lee, K. Y. S., Lui, A. S. Y., Kan, P. P. K., Luke, K. L., Mak, Y. M., Cheung, P. M., Cheng, L., & Wong, I. (2009). A case series on the social thinking training of mainstreamed secondary school students with high-functioning autism. *Hong Kong Journal of Mental Health, 35*, 10–17.

Leaf, J. B., Kassardjian, A., Oppenheim-Leaf, M. L., Cihon, J. H., Taubman, M., Leaf, R., & McEachin, J. (2016). Social thinking: science, pseudoscience, or antiscience? *Behavior Analysis in Practice, 9*(2), 152–157. https://doi.org/10.1007/s40617-016-0108-1.

Luiselli, J. K., Putnam, R. F., Handler, M. W., & Feinberg, A. B. (2005). Whole-school positive behavioral support: effects on student discipline problems an academic performance. *Educational Psychologist, 25*, 183–198. https://doi.org/10.1080/01443410420003012 65.

Maag, J. W. (2006). Social skills training for students with emotional and behavioral disorders: a review of reviews. *Behavioral Disorders, 32*(1), 4–17. https://doi.org/10.1177/019874290603200104.

March, R. E., & Horner, R. H. (2002). Feasibility and contributions of functional behavioral assessment in schools. *Journal of Emotional and Behavioral Disorders, 10*, 158–170. https://doi.org/10.1177/10634266020100030401.

Metcalf, T. (2015). *What's your plan? Accurate decision making within a multi-tier system of supports: critical areas in tier 2*. New York: RTI Action Network, A Program of the National Center for Learning Disabilities Retrieved from https://www.cnyric.org/tfiles/folder1306/Tier%202%20What's%20Your%20Plan.pdf.

Myles, B., Jones-Bock, S., & Simpson, R. (2001). *Asperger Syndrome Diagnostic Scale (ASDS)*. North Tonawanda: Multi-health systems (MHS).

Owens, J. S., & Murphy, C. E. (2004). Effectiveness research in the context of school-based mental health. *Clinical Child and Family Psychology Review, 7*(4), 195–209.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education Retrieved from http://ncer.ed.gov.

PBIS National Technical Assistance Center. (2009). *Is school-wide positive behavior support an evidence-based practice?* Washington, DC: Office of Special Education Programs (OSEP) Retrieved from http://www.pbis.org/research.

Ramdoss, S., Machalicek, W., Rispoli, M., Mulloy, A., Lang, R., & O'Reilly, M. (2012). Computer-based interventions to improve social and emotional skills in individuals with autism spectrum disorders: a systematic review. *Developmental Neurorehabilitation, 15*(2), 119–135. https://doi.org/10.3109/17518423.2011.651655.

Sansosti, F. J. (2010). Teaching social skills to children with autism spectrum disorders using tiers of support: a guide for school-based professionals. *Psychology in the Schools, 47*(3), 257–281. https://doi.org/10.1002/pits.20469.

Scott, T. M., McIntyre, J., Liaupsin, C., Nelson, C. M., Conroy, M., & Payne, L. (2005). An examination of the relation between functional behavior assessment and selected intervention strategies with school-based teams. *Journal of Positive Behavior Interventions, 7*, 205–215. https://doi.org/10.1177/10983007050070040201.

Shukla-Mehta, S., Miller, T., & Callahan, K. J. (2010). Evaluating the effectiveness of video instruction on social and communication

skills training for children with autism spectrum disorders: a review of the literature. *Focus on Autism and Other Developmental Disabilities, 25*(1), 23–36. https://doi.org/10.1177/1088357609352901.

Sugai, G. (2007). Promoting behavioral competence in schools: a commentary on exemplary practices. *Psychology in the Schools, 44*(1), 113–118. https://doi.org/10.1002/pits.20210.

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245–259.

Sugai, G., & Simonsen, B. (2012). *Positive behavioral interventions and supports: history, defining features, and misconceptions*. Storrs, CT: University of Connecticut, Center for PBIS & Center for Positive Behavioral Interventions and Supports Retrieved from http://idahotc.com/Portals/6/Docs/2015/Tier_1/articles/PBIS_history.features.misconceptions.pdf.

The Language Express. (2013). *The social express teaching manual* (1st ed.). San Diego: Author.

Winner, M. G. (2005). *Think social! A social thinking curriculum for school-age students*. San Jose: Think Social Publishing, Incorporated.

Winner, M. G., & Crooke, P. J. (2009). Social thinking: a training paradigm for professionals and treatment approach for individuals with social learning/social pragmatic challenges. *Perspectives on Language Learning and Education, 16*(2), 62–69. https://doi.org/10.1044/lle16.2.62.

S. Kathleen Krach is an Assistant Professor at Florida State University in the area of school psychology program. Her research interest is in cultural and linguistic assessment using technology.

Michael McCreery, Ph.D. is an Associate Professor in the Department of Teaching and Learning at UNLV. His research interests are in computer-basedtools for social, emotional, and behavioral assessments.

Kanessa M. Doss, Ph.D. is an Associate Professor in the Psychology Department at Troy University. Her research interests are in the field of diversity and school psychology.

Dasha Highsmith, Ed.S. is a doctoral candidate in the Combined School and Counseling Psychology at Florida State University.