Moderators of School Intervention Outcomes for Children with Autism Spectrum Disorder

Christopher Lopata[a]

James P Donnelly[a]

Marcus L Thomeer[a]

Jonathan D Rodgers[a]

Jennifer Lodi-Smith[a]

Adam J Booth[a]

Martin A Volker[b]

[a]Institute for Autism Research, Canisius College, Buffalo, NY, USA
[b]Center for Research in Autism, Intellectual, and Other Neurodevelopmental Disabilities, Michigan State University, East Lansing, MI, USA

Corresponding author: Christopher Lopata, Institute for Autism Research, Canisius College, Science Hall 1016C, 2001 Main Street, Buffalo, NY, 14208, USA; Ph. 716-888-2861; Email: lopatac@canisius.edu

Publication date: May 1, 2020

IES award number: R324A130216

Citation of Journal Publishing Article: Journal of Abnormal Child Psychology v48 n p1105-1114 May 2020

Moderators of School Intervention Outcomes for Children with Autism Spectrum Disorder

Abstract

A prior cluster randomized controlled trial (RCT) compared outcomes for a comprehensive school intervention (schoolMAX) to typical educational programming (services-as-usual [SAU]) for 103 children with autism spectrum disorder (ASD) without intellectual disability. The schoolMAX intervention was superior to SAU in improving social-cognitive understanding (emotion-recognition), social/social-communication skills, and ASD-related impairment (symptoms). In the current study, a range of demographic, clinical, and school variables were tested as potential moderators of treatment outcomes from the prior RCT. Moderation effects were not evident in demographics, child IQ, language, or ASD diagnostic symptoms, or school SES. Baseline externalizing symptoms moderated the outcome of social-cognitive understanding and adaptive skills moderated the outcome of ASD-related symptoms (no other comorbid symptoms or adaptive skills ratings moderated outcomes on the three measures). Overall, findings suggest that the main effects of treatment were, with two exceptions, unaffected by third variables.

Keywords: comprehensive school intervention, schoolMAX, ASD, moderators

Introduction

Children with autism spectrum disorder (ASD) without intellectual disability (ID) comprise more than two-thirds of those diagnosed (Christensen et al., 2016). Although these children exhibit relative cognitive and language strengths, their social deficits and restricted and repetitive behaviors significantly interfere with adaptive social functioning (American Psychiatric Association, 2013). Social impairments of these children include deficits in basic social behaviors/skills (e.g., initiating and responding to social bids) and more complex social-cognitive understanding (e.g., interpreting facial expressions and others' perspectives; Bellini, Gardner, & Markoff, 2014). Higher levels of stereotyped and repetitive behaviors/interests have also been associated with lower adaptive and social performance (McDonald et al., 2015). Comorbid externalizing and internalizing symptoms are also common and pose additional challenges. For example, disruptive externalizing behaviors can lead to more restrictive school placements and intensive treatments (Lecavalier et al., 2017) and internalizing symptoms (e.g., anxiety) can further interfere with social performance (White et al., 2013).

Addressing the core social impairment is essential and social interventions are commonly used to target characteristic impairments in social-cognition, social skills/behaviors, and ASD symptoms (Scarpa, White, & Attwood, 2013). Meta-analyses of social intervention RCTs in university/lab settings for youth with ASD without ID suggested moderate effects overall and variability in outcomes, highlighting the need for ongoing development of treatments and testing in RCTs (Gates, Kang, & Lerner, 2017; Reichow, Steiner, & Volkmar, 2012). These reviews suggest some promise for social interventions when administered in university/lab settings; however, the effects rarely transfer to more authentic settings including schools leading to calls for development and testing of social interventions within schools (Kasari et al., 2016;

McMahon, Lerner, & Britton, 2013). The challenges of developing effective interventions for authentic settings have long been recognized. Early reviews comparing the efficacy lab-based child therapy to therapy delivered in real-world clinic settings revealed that the overall effects for clinic-based studies were far lower (often negligible) than the moderate-to-large effects reported for lab-based trials (Weisz, Weiss, & Donenberg, 1992; Weisz, Donenberg, Weiss, & Han, 1995). Differences between lab and clinic studies (severity of child disturbance, use of behavioral treatments, resource availability) reportedly contributed to the disparity in efficacy. Given this gap between outcomes, the authors recommended that effective lab-based techniques be tested in authentic settings under real-world conditions.

Despite indications of efficacy, there has been limited testing of social interventions in schools for children with ASD (Kasari & Smith, 2013), and even when conducted, many have used study staff to implement the intervention (i.e., non-school staff; Camargo et al., 2014; Iadarola et al., 2018). Developing comprehensive interventions that address the range of impairments and are feasible and effective in schools is especially challenging given competing demands and priorities, time constraints, and staffing requirements (Kasari & Smith, 2013; Odom, Boyd, Hall, & Hume, 2014). The lack of school social interventions is especially problematic as schools are a primary source of intervention for many children with ASD (Iadarola et al., 2018) and they provide extensive practice opportunities in an authentic setting (Ho, Stephenson, & Carter, 2018). School social interventions may have additional benefits as they target development of social competencies in the environments in which the children are expected to use them (Kasari et al., 2016).

Lopata, Thomeer, and colleagues (2019) developed a comprehensive school intervention (schoolMAX) for elementary school children with ASD without ID and tested it in a cluster

RCT. The schoolMAX intervention was superior to typical educational services (services-as-usual; SAU) on several social performance and ASD symptom indicators. That study, which included 103 children, enrolled in 35 public schools (17 schools randomly allocated to the intervention and 18 schools to SAU), was the largest RCT of a school-staff delivered multi-component intervention. Results of the linear mixed effects model analyses found the schoolMAX intervention was superior to SAU on the primary measures of social-cognition (emotion recognition testing by masked evaluators; Cambridge Mindreading Face-Voice Battery for Children, d=1.41; p<.001) and ASD symptoms (parent-teacher ratings composite; Social Responsiveness Scale 2nd Edition, d=-1.15; p<.001) and secondary measure of social/social-communication skills (parent-teacher ratings composite; Adapted Skillstreaming Checklist, d=1.29, p=.001). No differences were found on recess social interactions or academic skills.

Although some evidence has supported the efficacy of social interventions for youth with ASD without ID, significant variability across studies (e.g., child characteristics, intervention features, outcome measures) has made it difficult to draw conclusions about their efficacy (McMahon et al., 2013). Variability in treatment responsiveness also indicates a need to determine which children benefit from an individual intervention (Reichow & Barton, 2014). Despite the need for such studies, there has been limited testing of moderators of treatment outcomes. Testing moderators is essential to determine for whom an intervention may be more or less effective (Hinshaw, 2007; Lecavalier et al., 2017). In a study illustrating the influence of moderators, Hinshaw (2007) explored whether a number of child and family variables moderated outcomes for a large multimodal treatment trial for children with ADHD. Results indicated that factor such as child sex and comorbid oppositional and conduct disorder did not moderate outcomes, whereas comorbid anxiety disorder, ADHD severity, child IQ, and SES did (although

the pattern differed based on treatment modality). Studies such as this can help identify individuals/subgroups and contexts for which a treatment is particularly effective.

To date, examination of moderators in ASD social intervention studies has been hindered by several factors. For example, testing potential moderators requires larger samples (Kasari & Smith, 2013; Odom et al., 2014) and many social intervention studies have used small samples including school studies (Kaat & Lecavalier, 2014; Kasari & Smith, 2013). Examination of moderators of outcomes across studies has also been hindered by incomplete reporting of sample characteristics (demographics, IQ, ASD severity, etc.; Ho et al., 2018). Although research is limited and findings are inconsistent, there is widespread recognition of the need to test the potential moderating effects of child characteristics (e.g., IQ, language ability, age, sex, ASD symptom severity, comorbid symptoms) on intervention outcomes (Gates et al., 2017; Ho et al., 2018; McMahon et al., 2013; Reichow et al., 2012). Some studies have suggested that IQ, language level, and/or age are potentially predictive of social intervention outcomes (e.g., Solomon, Goodlin-Jones, & Anders, 2004; Whalon, Conroy, Martinez, & Werch, 2015), whereas others have found no effects (Gates et al., 2017; Koenig et al., 2010). Other studies have suggested that comorbid symptoms may play a moderating role. Kasari et al. (2016) found baseline ratings of behavior problems moderated outcomes differently for two school social interventions (adult directed vs. peer-mediated) for children with ASD. They concluded that efficacy may differ based on the specific intervention and co-occurring child behaviors. A study comparing parent-delivered interventions for disruptive behaviors of young children with ASD (of variable IQ levels) found neither IQ nor ASD symptom severity moderated outcomes; however, baseline symptoms of ADHD, anxiety, and oppositional defiant disorder moderated outcomes for irritable and noncompliant behaviors (lower comorbid levels generally associated

with larger treatment effects; Lecavalier et al., 2017). SES is another potential moderator. In the study by Lecavalier et al. (2017), household income moderated the outcome of noncompliance (higher income associated with better outcomes). Economic and resource factors may also moderate school intervention outcomes. Lower SES schools (characterized by a higher percentage of free-or-reduced lunch; January, Casey, & Paulson, 2011) often struggle with intervention implementation due to limited resources (Iadarola et al., 2018). Considering these variables, there is a need to test the potential moderating effects of child features and context on intervention outcomes (Gates et al., 2017).

A final consideration involves efforts to test moderators across social intervention studies via meta-analyses. As illustrated in the meta-analysis by Gates et al. (2017), there is significant variability in social intervention studies (e.g., child characteristics, intervention features, outcome measures) and outcomes. Pooling data may mask potential moderators due to highly variable samples, interventions, methods, and measures across studies. Because of this, studies are needed of moderators of outcomes for individual social interventions for children with ASD without ID (Kaat & Lecavalier, 2014; McMahon et al., 2013). Additionally, social intervention studies for these children often include multiple measures of social performance (social knowledge and social skills/impairments; Ho et al., 2018; McMahon et al., 2013) and it is important to test potential moderators for the different outcome indicators.

This study examined the potential moderating effect of demographic, clinical, and school variables on the significant outcomes from the prior RCT of the schoolMAX intervention for children with ASD without ID. Moderator variables were selected *a priori* based on the social and behavioral intervention literature for youth with ASD. Identifying moderators of outcomes is a critical next step in social intervention research for children with ASD without ID and is

particularly important for comprehensive school interventions as these can be especially difficult for school staff to implement in authentic settings (Kasari & Smith, 2013).

Method

Design

The background, methods, and results from the prior RCT have been previously reported (Lopata et al., 2019). As such, only a brief overview is provided here. The prior trial was approved by the Institutional Review Board (Canisius College IRB) and written parental consent and child assent were obtained prior to inclusion. Children who met inclusion criteria were randomly assigned by school building (cluster) to the intervention or SAU; buildings were randomly assigned based on the group-delivered nature of some of the intervention components and need to avoid cross-condition contamination within a building. The randomization sequence was generated separately by the study's biostatistician and transferred to the study personnel (ensuring independence in the allocation process). Children allocated to the intervention received the manualized intervention and those allocated to SAU received their typical educational programming throughout the school year. Baseline assessments were conducted six weeks into the school year (prior to initiation of the intervention) and follow-up assessments two weeks prior to the end of the school year for children in both conditions. Assessments that yielded significant effects included a child test of social-cognition (CAM-C) and parent-teacher ratings of ASD-related symptoms (SRS-2) and social/social-communication skills (ASC). The social-cognitive testing (CAM-C) was conducted by independent blinded evaluators.

Participants

The sample was recruited from public elementary schools in mainly suburban districts. Eligibility criteria were a diagnosis of ASD (confirmed via the Autism Diagnostic Interview-

Revised [ADI-R]; Rutter, LeCouteur, & Lord, 2003), Wechsler Intelligence Scale for Children-4th Edition (WISC-IV; Wechsler, 2003) short-form IQ >70 (and short-form VCI or PRI ≥80), and Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999) short-form expressive or receptive language score ≥75. A total of 114 children were screened, with 103 children, ages 6-12 years (grades 1-5), from 35 schools meeting inclusion criteria and enrolling. Seventeen schools (52 children with ASD) were randomly assigned to the intervention and 18 schools (51 children with ASD) to SAU. One child withdrew from the SAU condition without explanation. Each school contained a mean of three (*SD*=1.2) child participants. Baseline equivalence was demonstrated in no significant differences between conditions on any demographic, clinical, school, or baseline outcome measure.

Outcome Measures

Cambridge Mindreading Face-Voice Battery for Children (CAM-C; Administered to Child). The CAM-C (Golan & Baron-Cohen, 2006) measures social-cognitive understanding/ skills involving emotion recognition for 15 emotion concepts using facial expression video clips and speech audio clips. Children view/listen to each clip on a computer and select 1 of 4 emotion words that reflects the emotion of the person in the clip; higher total scores indicate greater accuracy. Test-retest reliability over a 10-15 week interval was 0.74-0.76. The CAM-C accurately differentiates children with ASD without ID from typical children and its scores are negatively correlated with ASD symptoms (Golan, Sinai-Gavrilov, & Baron-Cohen, 2015). The CAM-C was administered by independent blinded evaluators.

Social Responsiveness Scale, 2nd Edition, School Age Form (SRS-2; Completed by Parents and Teachers). The SRS-2 (Constantino & Gruber, 2012) is an objective measure of ASD-associated impairments/symptoms. It consists of 65 items that assess ASD features

involving interpersonal behavior, communication, and stereotypic and circumscribed behaviors and interests on a continuous scale. Informants rate the frequency of behaviors on a scale of 1 (not true) to 4 (almost always true), with higher total scores indicating greater ASD-associated symptom severity/impairments. The total score has internal consistency estimates of 0.92-0.97 and the test accurately discriminates ASD and non-ASD samples. The mean of the parent-teacher ratings (*T*-scores) for each child was used in the analyses.

Adapted Skillstreaming Checklist (ASC; Completed by Parents and Teachers). The ASC (Lopata, Thomeer, Volker, Nida, & Lee, 2008) measures social/social-communication skills and behaviors of children with ASD without ID. Across the 38-items, 32 assess social/social-communication skills and 6 assess behavioral regulation and flexibility. Each item is rated on a scale of 1 (almost never) to 5 (almost always). Items are summed to yield a total composite and higher total scores indicate greater use of prosocial adaptive skills/behaviors. Studies of the ASC for children with ASD without ID (Lopata et al., 2017; Lopata et al., in press) yielded internal consistency estimates of 0.92 (parent ratings) and 0.93 (teacher ratings). Validity was supported by moderate-to-high inverse correlations with ratings of ASD symptoms and problem behaviors and positive correlations with prosocial/adaptive skills on established scales. The mean of the parent-teacher ratings (total scores) for each child was analyzed.

Moderator Variables

Demographic and setting variables. Five variables were tested including parent education, school SES and fidelity, and child age and sex. Parent education was based on the child's primary residence and was measured as the highest year of education completed for the individual biological parent in single-parent households (*n*=21, 20%) and mean of the highest year of education completed for married parents living in the same household (both biological

parents *n*=79, 77%; one biological parent and one step-parent *n*=2, 2%; adoptive parents *n*=1, 1%). School SES was measured based on the percentage of students within a building receiving free-or-reduced lunch. School implementation accuracy (fidelity) was measured throughout the school year by research assistants using standardized fidelity checklists. For the intervention group, accuracy was defined as the percentage of intervention elements delivered per the manualized protocol. Within the SAU schools, the same fidelity forms (with sequencing requirements removed) were completed to document any elements from the intervention that might have been present. Child age was measured in years at the time of study enrollment.

Child IQ, language ability, and ASD diagnostic symptoms. Overall IQ was measured using a 4-subtest short-form of the WISC-IV (Wechsler, 2003) consisting of the Block Design, Similarities, Vocabulary, and Matrix Reasoning subtests, with the short-form VCI calculated using the Similarities and Vocabulary subtests and PRI using the Block Design and Matrix Reasoning subtests. Language ability was assessed using a short-form of the CASL (Carrow-Woolfolk, 1999), with expressive language measured using the Antonyms and Syntax Construction subtests and receptive language using the Synonyms and Paragraph Comprehension subtests. ASD diagnostic symptoms were assessed using the ADI-R (Rutter et al., 2003) which is a standardized interview administered to a parent or caregiver familiar with the developmental history and current behavior of the child. It assesses symptoms in the areas of Reciprocal Social Interactions, Language/Communication, and Restricted, Repetitive and Stereotyped Patterns of Behavior/Interests. The WISC-IV, CASL, and ADI-R were administered at screening.

Child comorbid symptoms and adaptive skills. Comorbid symptoms and adaptive skills were assessed at baseline using the Behavior Assessment System for Children-Second Edition (BASC-2; Reynolds & Kamphaus, 2004), Parent Rating Scales (PRS) and Teacher

Rating Scales (TRS). The BASC-2 PRS and TRS measure clinical symptoms and adaptive skills to assist with diagnosis, as well as intervention planning. This study used the Child (6-to-11 years; PRS-C and TRS-C) and Adolescent (12-to-21 years; PRS-A and TRS-A) forms to assess externalizing and internalizing symptoms, and adaptive skills. The BASC-2 has consistent scales across age levels which, "provides a basis for consistent interpretation of scales" (Reynolds & Kamphaus, 2004, p. 2). The Externalizing composite is derived from scales measuring symptoms of hyperactivity, aggression, and conduct problems and the Internalizing composite is derived from scales measuring symptoms of anxiety, depression, and somatization; for both composites, higher scores (*T*-scores) indicate higher levels of comorbid symptoms. The Adaptive Skills composite is derived from scales measuring adaptability, social skills, leadership, communication, and performance of basic home/school tasks; lower scores (*T*-scores) on this composite indicate more severe deficits. Each item is rated on a 4-point frequency scale from 0 (never) to 3 (almost always). The BASC-2 PRS and TRS have strong internal consistency and correlate well with comparable scales on established measures of clinical symptoms and adaptive skills (Reynolds, & Kamphaus, 2004). Consistent with the outcome measures in the original RCT, the mean of the parent-teacher ratings (*T*-scores) for each child was used in the analyses.

Treatments

As noted, a detailed description of the protocol was previously published so only an overview is described here. schoolMAX is a manualized intervention delivered by school staff during the school year consisting of social skills groups (SSGs), emotion recognition instruction, therapeutic activities (TAs), a behavioral reinforcement system (Individual Daily Note [IDN]), and parent training (PT). The intervention uses cognitive and behavioral strategies to teach and reinforce social-cognitive understanding and social skills/behaviors. School staff were trained

during the summer and established fidelity with the protocol. The 30-hour manualized training included classroom instruction and applied practice exercises. During the applied exercises each staff member demonstrated ≥90% fidelity administering her/his component(s), assessed by research assistants using the standardized fidelity checklists. Implementation accuracy (fidelity) was defined as the percentage of intervention elements delivered per the protocol.

Different members of each student's educational team implemented a different component(s) (based on position/role, training, and/or experience) to avoid overburdening any individual. SSGs were conducted 2-3 times per week (total of 60-90 minutes) and each included up to 6 students with social impairments including 1-3 target students with ASD. Emotion recognition instruction was provided 2-3 times per week (total of 60 minutes) using the *Mind Reading* interactive software which teaches recognition of emotions in facial and vocal expressions (Baron-Cohen, Golan, Wheelwright, & Hill, 2004); participants received this instruction individually via computer. TAs were conducted 2 times per week (total of 40-60 minutes). These cooperative group activities included general and/or special education peers and were conducted to practice and reinforce targeted social and emotion recognition skills, and promote interest expansion. The IDN was implemented throughout the school day by all members of the student's educational team to practice and reinforce targeted skills and reduce ASD symptoms/problem behaviors. Each IDN included 3-5 operationally-defined targets and performance criteria, and students received feedback during and at the end of each day. Each student could earn 1 point per target per interval, with those earning ≥75% of the daily points receiving a home reward (reinforcer) provided by parents. PT was conducted monthly (60-90 minutes per session) during the school year by 1 or more members of the school team. PT sessions were manualized and targeted increased understanding of schoolMAX (content and

teaching procedures), home-school communication, and integration across settings, and ensured establishment of parent-provided reinforcers at home for school performance on the IDN. Overall fidelity was high across school buildings (*M*=96.0%; range 93.4%-98.8%).

Children in the SAU schools received their typical educational programming (including legally mandated services per their IEPs). Observations were conducted in the SAU schools by research assistants using the fidelity forms (with sequencing requirements removed) to identify any schoolMAX intervention elements that might have been provided. Results indicated that the schoolMAX intervention elements were rarely, if at all, observed (*M*=2.6%; range 1.0%-10.4%).

Data Analysis

The analysis was an exploratory assessment of possible moderation effects on the treatment outcomes of the prior trial. The original multilevel model that included school cluster as a random effect and treatment condition as a fixed effect was used as the base model to test the addition of interactions of potential moderators with treatment condition. All variables were examined for outliers, normal distributions, and missing data. Complete data were available for all cases. One participant obtained a WISC-IV VCI score of 135, three standard deviations above the sample mean. There were no other outlying scores for this child. Another participant obtained a score of 12 on the ADI-R Restricted and Repetitive Behavior scale, three standard deviations above the sample mean. No other outlying scores were evident in this case. Distributions of continuous moderators and outcomes were approximately normal. Continuous moderator candidates, including the WISC-IV, CASL, ADI-R, and BASC-2, were centered by standardizing with a mean of 0 and standard deviation of 1. The restricted range of the implementation fidelity scores eliminated the possibility of examining potential moderating effects of fidelity. Statistical significance was set at .05. Significant interaction tests were

followed by plots to aid interpretation. Stata 16 was used to test statistical models and generate graphs showing predicted means bounded by one *SD* intervals of the moderator with 95% confidence intervals. In the interest of thorough exploration and transparency, and in anticipation of the possibility of future meta-analyses, fixed effects of all interaction tests are reported with regression coefficients, standard errors, *p*-values, and 95% confidence intervals.

Power

The original study was designed as a cluster RCT with power of .84 based on the following estimates from prior studies by this group: *d*=.99, ICC=.23, schools=30, *N*=90, alpha=.025. The final data set included 103 participants in 35 schools with effect sizes of 1.41 and -1.15 on the primary outcomes (CAM-C and SRS-2, respectively). The original power analysis was focused on two primary hypotheses and did not assess power for interactions. However, it was possible to estimate the minimum detectable effect size for the moderator analysis with the *PowerUp!-Moderator* program (Dong, Kelcey, Spybrook, & Maynard, 2017). Input for the following analysis setup was derived from the prior study. In a two-level cluster random assignment model with a continuous moderator at level 1, 35 clusters with 3 children per cluster, an ICC of .20, alpha=.05 (two-tailed), proportion of participants randomized=.50, proportion of variance in level 1 outcomes explained by level 1 covariates=.20, and power=.80, the minimally detectable effect size was estimated to be .45 [95%CI=.13, .77]. If alpha is set to a more conservative .01 level, the minimally detectable effect size with 80% power is .55. Setting alpha to .001 produces a minimally detectable effect size of .68 with power remaining at 80%. For reference, the lower bound estimate of the smaller effect size in the prior trial was -.53. The adequate power to test meaningful effect sizes in this analysis is balanced against the large number of tests in this exploratory study.

Results

Table 1 provides descriptive statistics for the sample demographics and baseline characteristics included as moderators, as well as the outcome measures. Results of the moderator tests are shown in Table 2. There was no evidence of moderation by child, family, or school demographics, or child IQ, language ability, or ASD diagnostic symptoms.

Moderation effects were evident for two behavioral scales. Externalizing behaviors appear to have impacted the treatment such that children in the treatment group who were lower in externalizing behaviors at baseline experienced a larger improvement in social-cognitive understanding/skills (p=.01). **Fig. 1** Plot of Moderation of Treatment on CAM-C by BASC-2 Externalizing Behaviors (predicted mean + 95% CI) displays this result, with the BASC-2 Externalizing scale in standard deviation units. In **Fig. 2** Plot of Moderation of Treatment on SRS-2 by BASC-2 Adaptive Skills (predicted mean + 95% CI), children who were lowest in Adaptive Skills demonstrated a greater reduction in SRS-2 ASD symptoms (p=.003). The figure also suggests that children in both groups who scored highest in adaptive skills did not differ in symptom change. The slope of the SAU line suggests that improvement in symptom severity may be associated with better adaptive skills even in the untreated group, but the test for this coefficient was not significant (p=.06). Further inspection of the adaptive skills scores showed that the trend in the SAU slope was not due to outliers. There were no other significant interactions for the other BASC-2 scales.

Discussion

Findings have indicated that social interventions have positive effects on the social performance of children with ASD without ID (Gates et al., 2017), yet variability in outcomes within and across studies suggests that moderator variables may play a critical role in efficacy

(Reichow & Barton, 2014). Significant variability in participant characteristics, interventions (dosage, content, settings), and outcome measures also suggests that testing moderators across studies (e.g., via meta-analyses) may fail to detect the contribution of moderators for a specific intervention. It may be more informative to assess moderators for a specific intervention that has yielded positive outcomes for these children. There has been limited testing of moderators of intervention outcomes, findings have been mixed, and there is a need for exploratory studies of the potential moderating effect of demographic, school, and clinical variables on outcomes. Studies of school-staff delivered social interventions are even more scarce (Camargo et al., 2014; Kasari & Smith, 2013), and almost nothing is known about moderators of their outcomes.

This study explored the potential moderating role of a range of demographic, school, and clinical features on outcomes of a large RCT testing a comprehensive school intervention for children with ASD without ID. The study tested the potential moderating effects across three outcomes (measures) critical to the functioning and performance of these children (social-cognitive understanding, ASD impairments/symptoms, and social/social-communication skills). Results were consistent for the moderating effects of the demographic variables and child screening measures. Findings suggested no moderating effects of parent education, school SES, or child age, sex, IQ, language ability, or ASD diagnostic symptoms on the three outcome measures. These results are promising and suggest that the schoolMAX intervention may be effective across the broad range of parent, school, and child features and diagnostic variables tested. For example, some have noted that certain levels of IQ or language ability may be needed for an intervention to be effective (Ho et al., 2018). In this study, outcomes did not vary across the range of IQ (full scale >70; VCI or PRI $\geq$80) or language ability (receptive or expressive $\geq$75) of the sample. Baseline ASD diagnostic symptoms also did not moderate

outcomes. Further, the outcomes were unaffected by school SES which can potentially affect resource availability and intervention implementation. In fact, school implementation accuracy (fidelity) was consistently high. This is very promising and supports the feasibility and efficacy of the intervention in authentic school settings. However, potential moderating effects of fidelity could not be assessed in light of the high level of fidelity and limited variability across schools.

When testing the potential moderating effects of comorbid symptoms and adaptive behaviors, results suggested minimal influence on intervention outcomes. No moderating effects were found for internalizing symptoms on the three outcomes, externalizing symptoms on ASD-related symptoms or social/social-communication skills outcomes, or adaptive skills on social-cognitive understanding (emotion recognition) or social/social-communication skills outcomes. Findings did suggest that two baseline measures had a narrow and small moderating effect. Externalizing behaviors appeared to have impacted the treatment such that children in the treatment group who were lower in externalizing behaviors showed a larger improvement in social-cognitive understanding. Adaptive skills also had a moderating effect, with children initially rated lower in adaptive skills demonstrating a greater reduction in ASD symptoms.

Overall, the pattern of results suggests that outcomes were largely unaffected by baseline comorbid symptoms and adaptive skills. The two that were statistically significant had narrow (each affected only a single outcome measure) and small effects. Although this suggests the intervention will likely require minimal adaptation, the two features that were significant may have some implications. For example, because children higher in externalizing behaviors showed less improvement on one outcome, a behavioral contingency could be added to the instruction targeting social-cognition in order to improve task attention or completion. Although the children were continually monitored by school staff, these children might require closer

monitoring and increased prompting to ensure the externalizing behaviors are not interfering with learning. Again, the need for such adaptations appears small given that externalizing symptoms did not moderate the other two outcomes. The second feature involved adaptive skills for which children lower at baseline showed a greater reduction in ASD-related symptoms. It might be that these children have more severe impairments when beginning treatment and more room to improve. It is also possible that parents and teachers are better able to observe ASD-related impairments in children with more severe adaptive deficits and the same degree of change could appear larger for this subgroup. Similar to externalizing symptoms, the magnitude of moderation for adaptive skills was small and it did not moderate the other two outcomes.

Moderation analysis addresses the critical issue of what works for whom. Within the limits of this exploratory study, results suggested that the main effects of treatment were, with only two exceptions, unlikely to have been affected by third variables. This provides initial support for the generalizability of treatment outcomes across the broad set of variables tested. Only two variables suggested a potential moderating effect; however, these effects were narrow and most importantly the magnitudes of effect were small, again suggesting minimal influence.

This study addressed the critical need for testing of moderators of social intervention and particularly school social intervention outcomes; however, several limitations warrant mention. Although this was one of the largest school-staff delivered comprehensive school intervention studies for children with ASD without ID, the sample was nonetheless small for the number of comparisons in this study. Testing of moderators was not the primary purpose of the original RCT and therefore this study was exploratory in nature and the results require replication. Spybrook, Kelcey, and Dong (2016) recommended that future trials should anticipate moderator analysis in power calculations and have developed software to assist in the process (Dong et al.,

2017). The *PowerUp!-Moderator* program was used in the present study to conduct a retrospective power analysis. This showed adequate power to detect moderate effects, but the number of statistical tests was large due to the absence of prior theory and research to restrict the set of putative moderators, an issue that requires improvements in theory as well as synthesis of available exploratory findings. Additionally, only one of the outcome variables in this study was completed by naïve independent evaluators, with two of the measures completed by parents and teachers who were not naïve to treatment condition. Use of a parent-teacher rating mean for each child on the outcomes and moderator tests of comorbid symptoms and adaptive skills is a further limitation. This was done to be consistent with the original RCT and limit the number of comparisons; however, future studies might consider testing outcomes and moderators for each informant group separately (assuming a sufficiently large sample size). The study was also limited by the characteristics of the children (age, IQ, and language ability), families, and schools. As such, the findings may not generalize to children with ASD outside the inclusion parameters, and/or families and school districts that differ from those in the original study. The lack of variability in implementation fidelity also precluded testing of its potential moderation of outcomes and future studies with greater variability should test the influence of fidelity on outcomes. A final note of caution is warranted regarding generalization of the findings to other interventions. Social interventions and studies have differed significantly in terms of participants and intervention characteristics. As such, it will be important to test moderators of outcomes for individual interventions to ensure important moderators do not go undetected due to variability in samples and intervention components across studies. Meta-analytic testing of moderators may be more valuable once a sufficient number of replication studies are completed on an individual intervention or enough studies using comparable samples and methods are available for pooling.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. J. (2004). *Mind reading: The intervention guide to emotions*. London: Jessica Kingsley (www.jkp.com).

Bellini, S., Gardner, L., & Markoff, K. (2014). Social skill interventions. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders: Vol. 2. Assessment, interventions, and policy* (4th ed., pp. 887-906). Hoboken, NJ: John Wiley & Sons.

Camargo, S. P. H., Rispoli, M., Ganz, J., Hong, E. R., Davis, H., & Mason, R. (2014). A review of the quality of behaviorally-based intervention research to improve social interaction skills of children with ASD in inclusive settings. *Journal of Autism and Developmental Disorders, 44*, 2096-2116. doi: 10.1007/s10803-014-2060-7

Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language*. Circle Pines, MN: American Guidance Services.

Christensen, D. L., Baio, J., Braun, K. V. N., Bilder, D., Charles, J., Constantino, J. N., et al. (2016). Prevalence and characteristics of autism spectrum disorder among children aged 8 years — Autism and developmental disabilities monitoring network, 11 Sites, United States, 2012. *MMWR*, 65(No. SS-3)(No. SS-3),1–23.

Constantino, J. N., & Gruber, C. P. (2012). *Social Responsiveness Scale, Second Edition (SRS-2)*. Torrance, CA: Western Psychological Services.

Dong, N., Kelcey, B., Spybrook, J., & Maynard, R. A. (2017). *PowerUp!-Moderator: A tool for calculating statistical power and minimum detectable effect size of the moderator effects*

in cluster randomized trials (Version 1.08) [Software]. Available from http://www.causalevaluation.org/

Gates, J. A., Kang, E., & Lerner, M. D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clinical Psychology Review, 52*, 164-181. doi: 10.1016/j.cpr.2017.01.006

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology, 18*, 591-617. doi: 10.10170S0954579406060305

Golan, O., Sinai-Gavrilov, Y., & Baron-Cohen, S. (2015). The Cambridge Mindreading Face-Voice Battery for Children (CAM-C): Complex emotion recognition in children with and without autism spectrum conditions. *Molecular Autism,* 6:22. doi: 10.1186/s13229-015-0018-z

Hinshaw, S. P. (2007). Moderators and mediators of treatment outcome for youth with ADHD: Understanding for whom and how interventions work. *Journal of Pediatric Psychology, 32*, 664-675. doi: 10.1093/jpepsy/jsl055

Ho, B. P. V., Stephenson, J., & Carter, M. (2018). Cognitive-behavioral approaches for children with autism spectrum disorder: A trend analysis. *Research in Autism Spectrum Disorders, 45*, 27-41. doi: 10.1016/j.rasd.2017.10.003

Iadarola, S., Shih, W., Dean, M., Blanch, E., Harwood, R., Hetherington, S., et al. (2018). Implementing a manualized, classroom transition intervention for students with ASD in underresourced schools. *Behavior Modification, 42,* 126-147. doi:10.1177/0145445517711437

January, A. M., Casey, R. J., & Paulson, D. (2011). A meta-analysis of classroom-wide interventions to build social skills: Do they work? *School Psychology Review, 40*, 242-256.

Kaat, A. J., & Lecavalier, L. (2014). Group-based social skills treatment: A methodological review. *Research in Autism Spectrum Disorders, 8*, 15-24. doi: 10.1016/j.rasd.2013.10.007

Kasari, C., Dean, M., Kretzmann, M., Shih, W., Orlich, F., Whitney, R., et al. (2016). Children with autism spectrum disorder and social skills groups at school: A randomized trial comparing intervention approach and peer composition. *Journal of Child Psychology and Psychiatry, 57*, 171-179. doi: 10.1111/jcpp.12460

Kasari, C., & Smith, T. (2013). Interventions in schools for children with autism spectrum disorder: Methods and recommendations. *Autism, 17*, 254-267. doi: 10.1177/1362361312470496

Koenig, K., White, S. W., Pachler, M., Lau, M., Lewis, M., Klin, A., & Scahill, L. (2010). Promoting social skill development in children with pervasive developmental disorders: A feasibility and efficacy study. *Journal of Autism and Developmental Disorders, 40*, 1209-1218. doi: 10.1007/s10803-010-0979-x

Lecavalier, L., Smith, T., Johnson, C., Bearss, K., Swiezy, N., Aman, M. G., et al. (2017). Moderators of parent training for disruptive behaviors in young children with autism spectrum disorder. *Journal of Abnormal Child Psychology, 45*, 1235-1245. doi: 10.1007/s10802-016-0233-x

Lopata, C., Rodgers, J. D., Donnelly, J. P., Thomeer, M. L., McDonald, C. A., & Volker, M. A. (2017). Psychometric properties of the Adapted Skillstreaming Checklist for high-

functioning children with ASD. *Journal of Autism and Developmental Disorders, 47*, 2723-2732. doi: 10.1007/s10803-017-3189-y

Lopata, C., Thomeer, M. L., Volker, M. A., Nida, R. E., & Lee, G. K. (2008). Effectiveness of a manualized summer social treatment program for high-functioning children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*, 890-904. doi: 10.1007/s10803-007-0460-7

McDonald, C. A., Thomeer, M. L., Lopata, C., Fox, J. D., Donnelly, J. P., Tang, V., & Rodgers, J. D. (2015). VABS-2 ratings and predictors of adaptive behavior in children with HFASD. *Journal of Developmental and Physical Disabilities*, *27*, 235-247. doi: 10.1007/s10882-014-9411-3

McMahon, C. M., Lerner, M. D., & Britton, N. (2013). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: A review and looking to the future. *Adolescent Health, Medicine and Therapeutics, 4*, 23-38. doi: 10.2147/AHMT.S25402

Odom, S. L., Boyd, B. A., Hall, L. J., & Hume, K. A. (2014). Comprehensive treatment models for children and youth with autism spectrum disorders. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders: Vol. 2. Assessment, interventions, and policy* (4th ed., pp. 770-787). Hoboken, NJ: John Wiley & Sons.

Reichow, B., & Barton, E. E. (2014). Evidence-based psychosocial interventions for individuals with autism spectrum disorders. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders: Vol. 2.*

Assessment, interventions, and policy (4th ed., pp. 969-992). Hoboken, NJ: John Wiley & Sons.

Reichow, B., Steiner, A. M., & Volkmar, F. (2012). Social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, Issue 7. Art. No.: CD008511.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children* (2nd ed). Circle Pines, MN: AGS.

Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism Diagnostic Interview-Revised*. Los Angeles: Western Psychological Services.

Scarpa, A., White, S. W., & Attwood, T. (2013). *CBT for children and adolescents with high-functioning autism spectrum disorders*. New York, NY: Guilford.

Solomon, M., Goodlin-Jones, B. L., & Anders, T. F. (2004). A social adjustment enhancement intervention for high functioning autism, Asperger's syndrome, and pervasive development disorder NOS. *Journal of Autism and Developmental Disorders, 34*, 649-668. doi: 10.1007/s10803-004-5286-y

Spybrook, J., Kelcey, B., & Dong, N. (2016). Power for detecting treatment by moderator effects in two- and three-level cluster randomized trials. *Journal of Educational and Behavioral Statistics*, *41*, 605-627. doi: 10.3102%2F1076998616655442

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th ed.). San Antonio, TX: The Psychological Corporation.

Weisz, J. R., Donenberg, G. R., Weiss, B., & Han, S. S. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 688-701. doi: 10.1037/0022-006X.63.5.688

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47*, 1578-1585. doi: 10.1037/0003-066X.47.12.1578

Whalon, K. J., Conroy, M. A., Martinez, J. R., & Werch, B. L. (2015). School-based peer-related social competence interventions for children with autism spectrum disorder: A meta-analysis and descriptive review of single case research design studies. *Journal of Autism and Developmental Disorders, 45*, 1513-1531. doi: 10.1007/s10803-015-2373-1

White, S., Ollendick, T., Albano, A. M., Oswald, D., Johnson, C., Southam-Gerow, M. A., et al. (2013). Randomized controlled trial: Multimodal anxiety and social skill intervention for adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*, 382-394. doi: 10.1007/s10803-012-1577-x

Table 1

Descriptive Statistics for Moderators and Outcome Measures

Characteristic	Treatment (n = 52)	SAU (n = 51)	t / Fisher's exact (p)	Treatment (n = 52)	SAU (n = 50)
Moderator					
Parent Education: Mean (*SD*)	15.76 (2.08)	15.41 (2.14)	.84 (.41)		
School SES: Mean (*SD*)	26.81 (14.31)	32.33 (17.28)	1.77 (.08)		
Child: Mean (*SD*)					
Age	8.65 (1.29)	9.01 (1.45)	1.32 (.19)		
WISC-IV IQ	103.82 (12.94)	100.94 (14.84)	1.05 (.30)		
WISC-IV VCI	103.04 (14.39)	100.21 (14.07)	1.01 (.32)		
WISC-IV PRI	103.82 (15.82)	101.50 (16.59)	.73 (.47)		
CASL Expressive Language	98.04 (15.10)	95.11 (14.52)	1.02 (.32)		
CASL Receptive Language	103.84 (17.49)	100.19 (16.22)	1.10 (.27)		
ADI-R Social Interactions	18.31 (5.91)	18.67 (5.72)	.31 (.76)		
ADI-R Communication	14.52 (3.91)	15.20 (5.43)	.73 (.47)		
ADI-R Repetitive Behavior	6.10 (1.72)	5.90 (2.24)	.50 (.62)		
BASC-2 Externalizing	57.12 (10.98)	57.96 (11.39)	.38 (.70)		

BASC-2 Internalizing	56.81 (14.32)	57.57 (13.07)	.28 (.78)		
BASC-2 Adaptive Skills	34.69 (8.61)	35.12 (7.59)	.27 (.79)		
Sex (male): *n* (%)	47 (90.4)	47 (92.2)	.10 (1.0)		
Outcome: Mean (*SD*)	Baseline	Baseline		Follow-up	Follow-up
CAM-C	46.04 (12.92)	46.09 (11.70)	.03 (.98)	58.73 (14.60)	48.76 (12.94)
SRS-2 Parent-Teacher	71.93 (9.98)	71.48 (7.04)	.27 (.79)	64.84 (8.13)	69.72 (9.23)
ASC Parent-Teacher	104.73 (17.98)	107.40 (13.33)	.86 (.39)	112.20 (17.13)	108.71 (14.03)

Note. SAU=Services-As-Usual; WISC-IV=Wechsler Intelligence Scale for Children-4th Edition (short-form); VCI=Verbal Comprehension Index; PRI=Perceptual Reasoning Index; CASL=Comprehensive Assessment of Spoken Language (short-form); ADI-R=Autism Diagnostic Interview-Revised; BASC-2=Behavior Assessment System for Children-Second Edition; CAM-C=Cambridge Mindreading Face-Voice Battery for Children; SRS-2=Social Responsiveness Scale, 2nd Edition; ASC=Adapted Skillstreaming Checklist.
*$p<.05$

Table 2

Summary of Exploratory Analysis of Moderators of Treatment Outcomes (Fixed Effects of Interaction)

Moderator	Outcome											
	CAM-C				SRS-2				ASC			
	b	*SE*	*p*	*95% CI*	*b*	*SE*	*p*	*95% CI*	*b*	*SE*	*p*	*95% CI*
Parent Education	-.06	.59	.92	-1.23, 1.12	-.80	.51	.12	-1.82, .22	.07	1.12	.95	-2.15, 2.29
School SES	-.04	.11	.74	-.26, .19	.04	.07	.54	-.10, .19	.13	.15	.41	-.18, .43
Age	2.24	1.24	.07	-.21, 4.69	-.33	.80	.69	-1.92, 1.27	1.10	1.72	.53	-2.33, 4.53
Sex	-4.40	6.06	.47	-16.43, 7.62	-.86	3.88	.83	-8.54, 6.83	.73	8.36	.93	-15.84, 17.31
WISC-IV IQ	-.05	.13	.71	-.30, .20	.01	.08	.95	-.15, .16	.24	.17	.16	-.10, .57
WISC-IV VCI	-.02	.12	.88	-.26, .22	-.01	.08	.78	-.16, .14	.09	.17	.59	-.24, .42
WISC-IV PRI	-.07	.11	.53	-.28, .14	.00	.07	.99	-.13, .13	.21	.14	.15	-.08, .49
CASL Expressive	-.09	.12	.44	-.32, .14	.02	.07	.77	-.13, .17	.10	.16	.52	-.21, .42
CASL Receptive	-.11	.10	.29	-.31, .09	-.02	.06	.82	-.14, .11	.08	.14	.58	-.20, .36
ADI-R Social	.28	.29	.35	-.30, .86	.30	.19	.11	-.07, .67	.22	.40	.59	-.58, 1.01
ADI-R Communication	-.35	.22	.35	-1.09, .39	-.08	.14	.57	-.35, .19	.62	.52	.23	-.41, 1.66
ADI-R Repetitive	1.65	.87	.06	-.08, 3.38	-.15	.57	.79	-1.28, .98	-.85	1.22	.49	-3.27, 1.57
BASC-2 Externalizing	**-.48**	**.21**	**.01**	**-.90, -.07**	.18	.14	.20	-.10, .45	-.07	.30	.82	-.66, .53
BASC-2 Internalizing	.28	.17	.11	-.06, .62	.17	.11	.11	-.04, .38	.06	.23	.81	-.41, .52
BASC-2 Adaptive	.31	.29	.30	-.27, .88	**-.57**	**.19**	**.003**	**-.95, -.20**	-.73	.41	.08	-1.55, .09

Note. WISC-IV=Wechsler Intelligence Scale for Children-4th Edition (short-form); VCI=Verbal Comprehension Index; PRI=Perceptual Reasoning Index; CASL=Comprehensive Assessment of Spoken Language (short-form); ADI-R=Autism Diagnostic Interview-Revised; BASC-2=Behavior Assessment System for Children-Second Edition; CAM-C=Cambridge Mindreading Face-Voice Battery for Children; SRS-2=Social Responsiveness Scale, 2nd Edition; ASC=Adapted Skillstreaming Checklist.

Statistically significant values bolded and underlined.

Figure 1

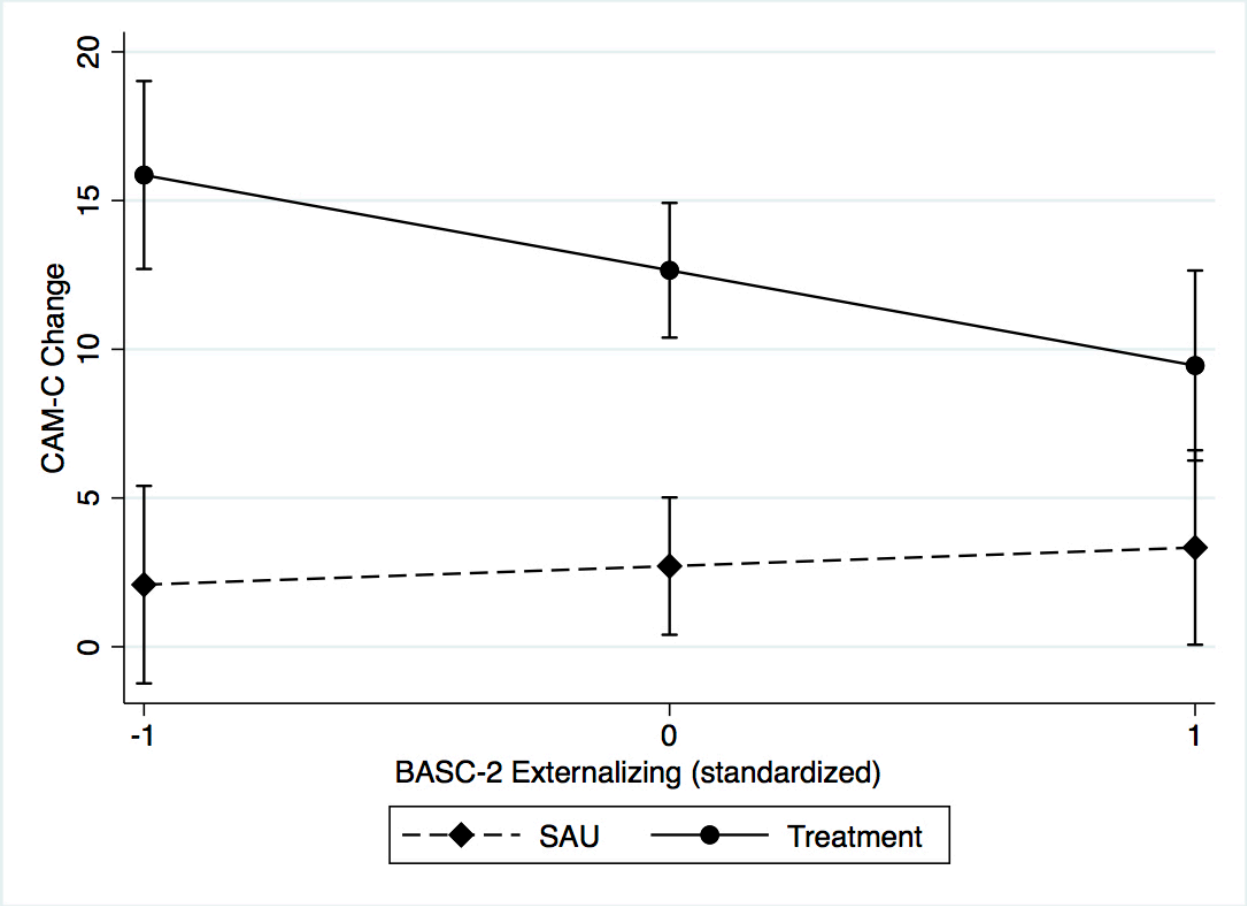

Figure 2