American Journal on Intellectual and Developmental Disabilities

Quality of Educational Programs for Elementary School-Age Students with Autism

--Manuscript Draft--

Manuscript Number:	AJIDD-D-20-00158R1
Article Type:	Research Report
Keywords:	Educational Programs; Quality; Autism
Corresponding Author:	Samuel L. Odom, Ph.D. University of North Carolina at Chapel Hill San Diego, CA UNITED STATES
First Author:	Samuel L. Odom, Ph.D.
Order of Authors:	Samuel L. Odom, Ph.D.
	Ann M. Sam, Ph.D.
	Brianne Tomaszewski, Ph.D.
	Ann W. Cox, Ph.D.
Manuscript Region of Origin:	UNITED STATES
Abstract:	The purpose of this study was to assess the quality of educational programs for school-aged children with autism in the United States. Investigators completed the Autism Program Environment Quality Rating Systems-Preschool/Elementary (APERS) in 60 elementary schools enrolling children with autism. The mean total rating scores were near the midpoint rating, indicating schools were providing educational program environments classified as adequate but not of high quality. Domains of the APERS reflecting structural quality tended to be significantly above average and domains reflecting process quality tended to be significantly below average. With a few exceptions, inclusive and special education program did not differ significantly in total program quality ratings and reflected this same pattern of domain quality ratings.

Quality of Educational Programs for Elementary School-Age Students with Autism

Abstract

The purpose of this study was to assess the quality of educational programs for school-aged children with autism in the United States. Investigators completed the Autism Program Environment Quality Rating Systems-Preschool/Elementary (APERS) in 60 elementary schools enrolling children with autism. The mean total rating scores were near the midpoint rating, indicating schools were providing educational program environments classified as adequate but not of high quality. Domains of the APERS reflecting structural quality tended to be significantly above average and domains reflecting process quality tended to be significantly below average. With a few exceptions, inclusive and special education program did not differ significantly in total program quality ratings and reflected the same pattern of domain quality ratings.

Quality of Educational Programs for Elementary School-Age Students with Autism

In the last 20 years, substantial increases in the number of children with autism attending public schools and receiving special education services parallels autism's[1] rapidly accelerating prevalence in society at large (U. S. Department of Education, 2019). Enrollment in public school programs brings an obligation for schools to provide an educational program of acceptable, and ideally high, quality. Yet to date, there is little information about the quality of educational programs for autistic children in elementary schools. The purposes of this study are to examine the overall quality of educational programs that elementary students with autism received as well as the quality of inclusive and primarily special education settings. In addition, the contextual factors associated with program quality will be examined.

Autism is a neurodevelopmental impairment, thought to be genetic in origin, and characterized by challenges in social communication and the presence of restricted interests and repetitive behavior (Jackson & Volkmar, 2019). Autism is a spectrum disorder, in that its social

[1] Currently there is disagreement about the terminology that authors should use as a descriptor for autism. The International Classification of Diseases-10 uses the term *autistic disorder*; the American Psychiatric Association (2013) uses the term *autism spectrum disorder* (ASD); and the Individuals with Disabilities Education Improvement Act of 2004 uses the term *autism.* The form of the terminology accepted primary by professional organizations is person-first, as in a child with ASD. Advocacy group and self-advocacy state that the term they prefer is identify-first, as in autistic child. In this paper and consistent with rising consensus in the field (Vivanti, 2020), the authors use all of these descriptors and forms to describe individuals with autism.

and behavioral challenges are expressed in a variety of ways with a wide range of support needed (i.e., from little to very substantial support) for activities of daily living and participation in school programs (American Psychiatric Association, 2013). Federal law mandates that public schools provide free and appropriate education for students with autism and other disabilities. The number of autistic children receiving special education services in elementary schools has increased 86% in the last decade (U. S. Department of Education, 2019). In 2017 (latest figures available), 702,742 children and youth with autism (i.e., from 3 to 22 years of age) received special education services. If they began their education at age 3, had perfect attendance, and left the school system at age 22, each student would have spent 1,101,600 minutes of their lives in the public-school system. Next to the time they spend with their families, autistic children spend more time in public education programs than any other setting. The quality of the public school experience has the potential for affecting the development and learning of children with autism and is a significant dimension of their overall quality of life.

In the field of education, efforts to assess comprehensive program quality has been sporadic. The most systematic efforts have focused on the quality of early childhood education [e.g., Early Childhood Environmental Rating Scale, 3rd Ed. (Harms et al., 2015), Class Assessment Scoring System (Pianta et al., 2008), preschool inclusion (*Inclusive Classroom Profile* (Soukakou, 2016), and after-school/out of school programs (Kuperminc et al., 2016)]. For early childhood education, the National Institute of Child and Human Development (2006) noted that program environment assessments capture both structural (e.g., features that respond to state regulations such as class size and organization) and process (i.e., features that focus on instruction and skills development) quality. For children with autism, state agencies have developed measures that document necessary program features (Crimmins et al., 2001; Liberta et

al., 2001), but to date, there has been no published evaluations of the quality of programs for elementary school-aged children with autism.

To assess the quality of the program environments for children with autism in public school settings, investigators with the National Professional Development Center on Autism Spectrum Disorder (NPDC) developed the *Autism Program Environment Rating Scale* (APERS; Odom et al., 2018). NPDC Investigators proposed that program quality provides the foundation on which to base the individualized delivery of evidence-based practices for students with autism (Odom et al., 2018). For school settings, Odom et al. (2012) indicated that program quality functions very much the way "common factors" function in clinical psychology (Deegear & Lawson, 2003). The assumption in clinical psychology is that common factors (e.g., alliance, empathy, expectations, cultural adaptation) need to be in place for individualized therapy to be effectively implemented (Wampold, 2015).

The construct of program quality is built through the assessment of specific features of school programs (i.e., domains). As can be seen from Figure 1, a set of domains that assess different classroom ecological features, interdisciplinary teaming, and family involvement merge to provide a measure of overall program quality. As noted previously, the domains consist of individual items scaled on a 1 (lowest) – 5 (highest) rating system, with item anchors for the 1, 3, and 5 ratings. The scale was constructed so that the 3 rating represented acceptable but not high quality. The definition of the individual domains and number of items in each is provided in the Appendix. To examine the reliability of the APERS, Odom et al. (2018) analyzed internal consistency from two independent data sets, revealing Cronbach alpha coefficients of .94-.96, respectively for total ratings and .60-.92 for individual domains (i.e., most were .75 or above). In addition, to establish construct validity the researchers conducted a confirmatory factors analysis

with an initial set of data, founds that all domains loaded significantly onto a single factors, and replicated with finding with a second independent data set.

The APERS has been used to assess quality in two previous studies in the United States. To evaluate the impact of the NPDC model for program improvement, investigators conducted the APERS at pretest and posttest for preschool, elementary, and high school programs in the United States, finding significant increases in total and domain quality ratings across the school year (Odom et al., 2013). To assess the quality of secondary education programs for adolescents with autism in the United States, Kraemer et al. (2019) collected APERS on 60 high school programs distributed equally among sites in California, North Carolina, and Wisconsin. The research team, which was not working directly with the schools, conducted the assessments in the first two months of the school year. Overall, schools had an average total score on the APERS slightly above the mid-level, with relative strengths in program environment, learning climate, team collaboration, and family communication, all of which represent a form of structural quality. For the process quality domains (e.g., social, communication, independence, behavior, transition, etc.), average scores were substantially below the adequate level. Also, the quality of the transition domain was significantly different (poorer) for students in the inclusive programs as compared with students in the special education program. Last, Kraemer at al. (2019) examined the influence of contextual variables (e.g., SES of students, urbanicity), finding community context (i.e., suburban location) was associated with total quality. These findings provided a glimpse into the quality of secondary education programs for autistic adolescent students in the U.S. and had implications for program improvement. As yet, the quality of program environments for elementary school-aged autistic children has not been conducted.

The purpose of this study was to examine program quality for autistic students enrolled in public elementary schools. The specific research questions are: (1) What is the total quality of educational programs for elementary school-age student with autism? (2) Are there relatively strong and challenging features of program quality as reflected by domains? (3) Do inclusive and special education programs differ in program quality? (4) Are there school or community contextual factors that affect program quality?

Methods

This study was part of a larger RCT that examined the efficacy of an educational model for students with autism (Sam et al., 2020). All data were collected at the beginning of the year before intervention procedures were implemented. The research methods were approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Settings

The study took place in 60 elementary schools located in a southeastern part of the United States. The schools contained 56 inclusive (IN) and 58 special education (SE) programs (i.e., not all schools had both types of programs). In IN programs, students with autism spent the majority of their school day in general education classrooms, usually with support services provided by a resource special education teacher. In SE programs, autistic students spent the majority of their school day in a special education class, usually with some opportunities during the day to participate in classes or activities out of the classroom (e.g., music, art, recess). The elementary schools contained kindergarten to 5th grade classes. All participating schools were publicly funded and had general education classes as well as special education classes (i.e., no schools only for students with disabilities or charter schools). At the district level, administrators approved participation in the study before schools were recruited. At the school-level,

participation was voluntary with school administrators and at least three key members of school staff in each school agreeing to participate before recruitment began.

Table 1 provides demographic information about the schools. Schools were located in rural/towns (n=15), suburban areas (n=20), and in cities (n=27). The average size of the schools was 610 students. Around 70% were Title 1 schools and on average 54% of students in the schools qualified for free and reduced lunch (FARL). In addition, school records provided information on the race/ethnicity classification of students in the school.

Generalizer

As noted, the entire sample of schools was in one southeastern state. Recruitment focused on schools in rural, suburban, and urban areas that had a mix of socioeconomic status and race/ethnicities of students, with the purpose of having a sample that might approximate demographics of the United States (Tipton, 2014). In order to assess the degree to which the schools in this sample were representative of schools around the country, we employed the Generalizer Index (Tipton & Miller, 2016). The generalizability index assesses the degree to which a sample is representative of an inference population (Tipton, 2014), which for this study is the United States. Scores range between 0 and 1 and are categorized as *very high* for scores between 1 and .90, *high* for scores between .90 and .80, *medium* for scores between .80 and .50, and *low* for scores below .50 (Tipton & Miller, 2016). The generalizability index was .92 representing very high generalizability.

Participants

Although the school program was the focus on the APERS assessment, students with autism who were participating in the larger study (an average of 8 per school) served as the reference point for observations (i.e., the classes in which they were enrolled and the activities in

which they engaged were the locations for the observations and guided to some extent the interview questions, to be described). All students had a primary or secondary educational diagnosis of autism, which qualified them for special education services in the southeastern state. The exception was a small number of students classified as having developmental delay but also had a clinical diagnosis of autism. Students' mean age was 7.83 years (SD= 1.81 years); 79% of the students were boys; and 43% of the students were white, nonHispanic (i.e., 57% of students came from a range of races and ethnicities). As can be seen in Table 2, autistic student participants in the study were evenly distributed across grade levels.

Measure: Autism Program Environment Rating Scale

For this study, the Autism Program Environment Rating Scale—Preschool/Elementary [referred to as the APERS (Odom et al., in press)] assesses the quality of programs serving learners with ASD who enrolled in Grades Preschool to Grade 5. The APERS-PE consists of 59 items organized into 10 domains (learning environment, positive learning climate, assessment and IEP development, curriculum and instruction, communication, social competence, personal independence and competence, functional behavior, family involvement, and teaming). Items are based on a five-point Likert-type rating continuum, with behavioral anchors at for the 1, 3, and 5 ratings. The "1" rating indicates the poor quality, the "3" rating indicates acceptable quality, and the "5" rating represents excellent quality. The coding is completed electronically, and the program calculates the mean rating for total and domain rating scores.

APERS Data Collection

Information for the APERS was gathered through observations, interviews, and document analysis. For the observations, a single independent rater (i.e., not otherwise involved with the school) observed for a minimum of six hours throughout a school day for schools that had IN

and SE programs (i.e., three hours for schools having only one type of program), sampling a representative variety of activities that typically occur (i.e., classroom instruction, transitions, lunch, recess, music or art). In some instances, observations were spread over multiple days (e.g., when an unusual event occurred on a day or the lead teacher absent). Up to three autistic students were selected by school staff as representative of students in the school and program. Identification of these "focal" children helped assessors plan the classes and activities to be observed, but other autistic students located in those contexts were also observed throughout the observational period.

The APERS assessor also interviewed key school staff members and parents, because some features of program quality could not be directly observed during the assessment period (e.g., team meetings, parent-teacher interactions, speech therapy sessions). Special education teachers (SE and IN resource teachers), general education teachers, related service personnel, the principal or other key member of school administration, and parents all participated in interviews. The APERS assessor followed standard protocols when conducting the interviews.

The IEPs of the three focal students were the primary data sources for document analysis. The IEPs contains the educational goals the students, which also provide a focus for the observations (i.e., to determine if instruction was addressing individual goals). In addition, other documents, such as behavior intervention plans, or relevant descriptions of the school were examined if relevant for any of the APERS-PE items. When all observations, interviews, and document analyses had been obtained, the APERS assessor reviewed all data that were collected. He/she then completed separate APERS ratings for the IN and SE programs.

Data Collector Training. All research staff participated in an APERS training. The APERS training included an eight-hour didactic training with a trained APERS assessor. The

didactic training consisted of an introduction to APERS domains and items, instructions on scoring with examples, and procedures for observations, interviews, and document analysis. APERS assessors then conducted a complete APERS assessment with a trained rater, examined agreement between ratings, and reached consensus on items on which they did not have exact agreement. Once training was complete, the assessor then completed APERS independently, which was reviewed by the trained APERS coder with an additional debrief between the trainer and the trainee. When the debrief indicated that the data collector had accurately followed protocols and ratings, they were then authorized to begin data collection.

Data Analysis

All analyses were performed in SAS 9.4. A repeated-measures ANOVA was performed utilizing data from 60 elementary schools to identify strengths and challenges within schools across the APERS domains. To examine predictors of school quality, a multiple regression was performed to assess the extent to which school characteristics (i.e., % of students receiving free and reduced lunch, % White students, Title 1 eligibility status of schools, and urbanicity) predicated overall school quality as measured by the APERS total score. Title 1 eligibility was coded as (0) Not Title 1 eligible or (1) Title 1 Eligible. Urbanicity was dummy coded such that effects for city and suburban schools were compared to those of rural schools. Finally, hierarchical linear models were performed to examine program differences between inclusive and separate special education school programs.

Reliability and Inter-rater agreement

Previous studies of the APERS have documented its reliability and validity with a broader sample of participants (Odom et al., 2018). To document reliability for the current study, we examined the internal consistency of and interrater agreement on the measure.

Cronbach alphas were computed for the total score and the domains for IN and SE settings separately (See Table 3). There were high alphas for the overall totals (.93 and .96), and the weighted total (.95). Alphas for the domains ranged from .56 to .89. To assess inter-rater agreement, a second member of the research team collected the APERS-PE simultaneously with the primary rater for 16.7 % of the sample. Inter-rater agreement was calculated at the item level. Average inter-rater agreement across items for which. there was an exact match and agreement within one point were 65.9% and 87% respectively. ICCs between the two raters also were calculated, yielding .97 and .98 coefficients for the total scores for IN and SE programs, respectively, and a range of .53 to .95 for individual domains. At the item level, the mean difference between the two raters (i.e., primary rater scored a 3 and second rater scored a 2, yield a one-point difference on the item) was .12 (sd = .24).

Results

Program Quality in Elementary Schools

Overall, the total quality of program environments for the schools in the current study was slightly above the acceptable levels identified in the APERS assessment (See Figure 2). There were variations in the mean item ratings across domains, with some domain mean item ratings relatively high quality and some domains relatively low reflecting quality challenges. To examine this variation, we conducted one-sample t-tests, used the midpoint range of the APERS (i.e., 3.0, which reflected acceptable quality) as the point of comparison, and then analyzed the differences of each domain mean item rating from the midpoint rating of 3.0. The standard deviation for the total score mean item rating was used in this analysis. Bonferroni adjustments ($p < .005$) were made to correct for multiple univariate tests. Mean item ratings for Learning Environment ($t(59) = 4.49$, $p < .001$, Cohen's $d = .66$), Social Climate ($t(59) = 6.38$, $p < .001$,

Cohen's *d* = .95), and Family Involvement (t(59) = 6.90 , p < .001, Cohen's *d* = 1.06), were significantly higher than the midpoint on the APERS-PE. Alternatively, mean item ratings for Communication (t(59) = -5.03 , p <.001, Cohen's *d* = -.73), Social (t(59) = -6.48 , p < .001, Cohen's *d* = -.96), Independence (t(59) = -6.39 , p <.001, Cohen's *d* = -.94), and Functional Behavior (t(59) = -3.11 , p = .003, Cohen's *d* = -.48), were significant below the acceptable quality midpoint.

Inclusive and Separate Program Differences

For the 56 inclusive programs and 58 separate programs, APERS-PE total and domain scores appear in Figure 3. We conducted hierarchical linear models to examine program differences accounting for the school programs nested within schools for each APERS domain and the APERS Total score. The total APERS-PE mean item rating was not significantly different for the two programs, although the Special Education programs rating was slightly below the 3.0 marker. For the domains, the inclusive programs had significantly higher environment, t (53) = 4.69, *p* < .001 Cohen's *d* =.76, and social, t (53) = 3.15, *p* = .003 Cohen's *d* = .32, school quality ratings than SE programs. SE programs had significantly higher assessment school quality ratings than inclusive programs, t (53) = 2.10, *p* = .04 , Cohen's *d* = .25.

Predictors of School Quality

To examine the school and community context variables that could affect program quality, the weighted APERS-PE total mean item rating was regressed onto percentage of students on FARL, percentage of white students (i.e., the proxy for race and ethnicity in the school), Title 1 eligibility of schools, and urbanicity (See Table 4). Overall, the model was significant F (5,54) = 2.80, *p* = .03, R^2 = .21, adj R^2 = .13. The only significant predictor of the model was the percentage of students on FARL, B = -.01, SE = .004, β = -.55, *p* = .02. To probe

this relationship further, we examined the Pearson correlations between FARL, Title 1 school status ($r = .75$, $p<.001$), and percentage of white students ($r = -.70$, $p < .001$) and Title 1 status and percentage of white students ($r = -.60$, $p<.001$), finding substantial associations among the three. This does suggest that FARL might well be serving as the proxy variable for SES in this sample. There was not an association between FARL and urbanicity.

Discussion

The purpose of this study was to examine the quality of elementary school programs for students with autism and is the largest study, to date, to address this issue. Regarding the first research question, on average the quality of elementary school programs was a little above the midpoint of the rating scale. This finding suggests that, in general, schools in this sample are providing programs of adequate quality for students with autism. Probing further in the second research question, schools appeared to be relatively stronger in the structural quality of programs they provide. Two domains, Learning Environment and Social Climate, were significantly greater than the APERS quality midpoint (3.0) marker. Learning Environment reflects the safety and organization of the program (e.g., no hazardous areas, established schedule and routines, established work areas, etc.), which is especially important for autistic students (O'Nions et al., 2018). Positive Social Climate refers to teachers' use of more positive than negative statements, positive recognition of students work, social greetings when they arrive, which should be a feature of any classroom setting. In addition, Family Participation also was a strength, which refers to teachers' communications with families and their involvement of family members in the child's program. These features of programs are aligned with both the requirements and the spirit of IDEA, as well as state regulations for buildings and classrooms and, as such, reflect the concept of structural quality. Two other structural features of programs are Teaming (i.e., IDEA

requires a multidisciplinary team being involved in students educational program planning) and Assessment and IEP Development (i.e., IDEA require the development of an Individualized Education Plan based on assessment information). Although neither of these domains on the APERS could be viewed as particular strengths or weaknesses, they both were in the range of acceptable quality.

Conversely, process quality (i.e., focus on interventions areas that match learning needs) appeared to be less than adequate. The mean ratings for the Social, Communication, Functional Behavior, and Independence domains were significantly below the benchmark rating that designated adequate quality. This is particularly troubling, because these skills areas are likely to appear often on IEPs of students with autism (Ruble & McGrew, 2013). The Instruction domain was slightly above the 3.0 criterion, suggesting that general instructional practices (e.g., clear delivery of instruction; instructional materials appropriate) were adequate, but it appears that they may not have been applied in the need areas just mentioned.

This pattern of findings is similar to the results from the Kraemer et al. (2019) study conducted in autistic adolescents' high school programs located in three different states. The more structurally-oriented features of quality tended to be higher in that study, while the more process-oriented features were significantly lower. Similar findings have also occurred in other disciplines. In studies of early childhood education in which researchers use the CLASS to measure quality in early childhood education programs, classrooms often have lower scores on their instruction variable, as compared to the two subtests that focus on classroom organization and emotional support (Hamre et al., 2014; Hartman et al., 2916; Hu et al., 2016).

Regarding the third research question, the IDEA stipulates that students with disabilities receive a free and appropriate education in the least restrictive environment (LRE). Historically

and even in contemporary discourse (Kaufman & Badar, 2016; Simpson & Sasso, 1991), there have been questions about and criticisms of inclusive programs as appropriate options for students with autism, with a consistent call for more research on the issue. In the current study, the total mean quality of IN and SE programs was not significantly different, although each program had specific relative domain strengths. It is important to note that the characteristics of the children in the two types of programs were different, with students in SE needing greater amounts of support and the students in the IN appearing to need less support and to be better able to access the general education curriculum. One interpretation of these data is that the superiority or inferiority of SE or IN is not supported by the findings of this study. The discussion of IN as the best or worst option for an individual autistic student might be reframed to center on matching the learning needs of the individual student with the capacity of the program environment to provide appropriate learning experiences given the amount of support needed.

Community contextual variables may influence school program quality for students with autism, and in this study, such contextual variables were represented by school demographics and urbanicity. In their study of secondary school programs, Kraemer et al. (2019) found that urbanicity had a significant influence on program quality. In the current study, urbanicity was not a significant predictor of quality, but FARL, a school-level proxy measure for SES, was significantly and inversely related to program quality. The other measures associated with SES (i.e., level of racial and ethnic diversity in the school as reflected by percentage white, Title 1 school status) were not significant predictors in the model. It is possible that when the FARL variable entered the model, it left little variance that could be explain by other variables associated with SES. The high correlations between FARL and both Title 1 eligibility and percentage of white students suggests just such a relationship. In this study, we chose to not

create a composite measure because the variables consist of different types of data (e.g., continuous, categorical), but the high correlations among variables do suggest that SES is the construct driving this finding.

The findings of this study have a variety of practical implications. First, it appears that while often addressing structural quality of programs adequately or well, the process quality may be a significant challenge for schools. As noted, teachers often report that they do not feel prepared to work with students having autism or other developmental disabilities (Knight et al., 2019). The process quality of programs for children and youth with autism does, however, appears to be a malleable factor that may be fostered by professional development. In their work in nine states, Odom et al. (2013) found that process quality improved over a one year time period in programs for children and youth with autism whose staff participated in a professional development and coaching that focused on quality improvement and teachers' use of EBPs. Working in high schools, Hume et al. (2021) implemented a comprehensive program through professional development and coaching, which resulted in significant increases in total school quality as process quality domains. As such, educational agencies that find low levels of quality in their programs for children with autism may consider implementing a professional development and coaching program that focuses on both program quality and teachers' use of evidence-based practices. In addition, family members, while appearing to receive good communication from schools and involvement in the IEP process in this study, should be vigilant about making sure the high quality intervention practices that relate most to their children's skill needs are being implemented in school programs.

As noted, in the current study, the program quality for student in IN and SE programs was similar, with each showing relative strengths. The one exception could be related to the

Social Competence domain, for which the quality was poor for both program types, but "less poor" for the IN programs. Certainly, curriculum and intervention practices exist for promoting social competence (Chester, Richidale, & McGillivrary, 2019), social engagement with peers (Odom, 2019), and development of social networks (Shih et al., 2019). For SE programs, this requires perhaps a greater effort because the students with autism are in separate classes without neurotypical peers, but there are successful examples of such opportunity occurring in SE settings (Carter et al., 2019).

Because SES at the school level appears to be associated with the quality of programs for students with autism, particular attention may need to be placed on monitoring and supporting quality improvement in such schools. This is important for several reasons. School funding resources may well be limited for programs with autistic students from low SES families, so there may be fewer resources for professional development and program improvement. Parents of these student with autism may have less personal and political capital to advocate for supplemental resources that might be needed. But perhaps most important, the proportion of autistic students from nonwhite racial/ethnic/linguistic groups was positively correlated with school FARL. For autistic students racial/ethnic/linguistic from families of color, this suggests that poorer quality related to the autism-specific features of the school program may be further compounded by resources and service disparities that generally exist for all students in schools with a low SES student body.

It is important to acknowledge several limitations of this study. First, the sample of schools were from only one state. The Generalizer did suggest that the demographic characteristics of the sample were representative of schools in the U. S. However, funding structures, curriculum, diagnostic and eligibility decisions, and other factors influencing school

programs are often decided at the state and local levels and may not be sufficiently reflected just be demographic representativeness. It is important to note that the current study replicated many of the findings from the Kraemer et al. (2019) study, which was conducted in high schools in three regionally separate states, and which varied on educational policy issues (e.g., inclusion/special education assignment of students). Nevertheless, replications of this study in other states would be needed to establish the further generality of these findings.

Second, although internal consistency and most measures of interrater agreement were strong, inter-rater agreement was only collected for 16% of the APERS-PE assessments. The study would have been strengthened if a higher proportion of the sample had included a second rater. Third, in this study, assessors who conducted the APERS assessment were not blind raters and so were aware that students were enrolled in inclusive or special education programs. In fact, such "blinding" would not have been possible, but nevertheless awareness of program enrollment could have created bias in ratings of the respective programs.

In conclusion, this study provided information about the quality of programs for autistic students enrolled in elementary schools. Overall, the quality of programs was in the acceptable range, with strengths of the programs appearing in the domains reflecting structural quality. Alternatively, the challenges and potential areas for program improvement in the future were related to process quality, that is those areas that address instruction and interventions related to high priority learning needs for many students with autism. The pattern of results, with a few exceptions, occurred for both IN and SE programs.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, Fifth Edition.* Arlington, VA: American Psychiatric Association.

Carter, E. W., Steinbrenner, J. D., & Hall, L J. (2019). Exploring feasibility and fit: Peer-mediated interventions for high school students with autism spectrum disorder. *School Psychology Review, 48* (2), 157-169. https://doi.org/10.17105

Chester, M., Richdale, A.L. & McGillivray, J. Group-based social skills training with play for children on the autism spectrum. *Journal of Autism and Developmental Disorders, 49,* 2231–2242. https://doi.org/10.1007/s10803-019-03892-7

Crimmins, D. B., Durand, V. M., Theurer-Kaufman, K., & Everett, J. (2001). *Autism Program Quality Indicators*. Albany: New York State Education Department.

Deegear, J., & Lawson, D. M. (2003). The utility of empirically supported treatments. *Professional Psychology: Research and Practice, 34*(3), 271–277. https://doi.org/10.1037/0735-7028.34.3.271

Hamre, B., Hatfield, B., Pianta, R., & Jamil., F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children's development. *Child Development, 85*(3), 1257-1274. https://doi.org//10.1111/cdev.12184

Harms, T., Clifford, R., & Cryer, D. (2015). *Early Childhood Environment Rating Scale (3rd ed.).* New York: Teachers College Press.

Hartman, S. C., Warash, B. G., Curtis, R., & Hirst, J. D. (2016). Level of structural quality and process quality in rural preschool classrooms. *Early Child Development and Care, 186*(12),1952-1960. http://doi.10.1080/03004430.2015.1137563

Hu, B. Y., Fan, X., Gu, C., & Yang, N. (2016) Applicability of the Classroom Assessment Scoring System in Chinese preschools based on psychometric evidence. *Early Education and Development, 27*(5), 714-734. https://doi.10.1080/10409289.2016.1113069

Hume, K. A., Odom, S. L., Duwalt, L., Hall, L. J., Kraemer, B., Steinbrenner, J., Bolt, D., & Tomaszewski, B. (2021). *Efficacy of a high school programs for students with autism.* Manuscript under review.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)

Jackson, S., & Volkmar, F. (2019). Diagnosis and definition of autism and other pervasive developmental disorders. In F. Volkmar (Ed.), *Autism and pervasive developmental disorders* (pp. 1-24). Cambridge, England: Cambridge University Press. https://doi.10,1917.9781108297769.002

Kauffman, J. M., & Badar J. (2016). It's instruction over place — not the other way around! *Phi Delta Kappan, 98*(4), 55-59. https://doi:10.1177/0031721716681778

Knight, V. F., Huber, N. B., Kuntz, E. M., Carter, E. W., & Juarez, A. P. (2019). Instructional practices, priorities, and preparedness for educating students with autism and intellectual disability. *Focus on Autism and Other Developmental Disabilities, 34*(1), 3-14. https://doi.110.1177/10883576198812177

Kraemer, B. R., Odom, S. L., Tomaszewski, B., Hall, L. J., DeWalt, L., Hume, K. A., Steinbrenner, J. R., Szidon, K., & Brum, C. (2019). Quality of high school programs for students with autism spectrum disorder. *Autism, published prepublication online.* https://doi:10.1177/1362361319887280

Kuperminc, G. P., Seitz, S., Joseph, H., Khatib, N., Wilson, C., Collins, K., & Guecsnour, O. (2019). Enhancing program quality in a national sample of after-school settings: The role

of youth-staff interactions and staff/organizational functioning. *American Journal of Community Psychology, 63,* 391-404. https://*doi:10.1002/ajcp.12329*

National Institute of Child Health and Human Development. (2006).Study of early child care and youth development (SECCYD): Findings for children up to Age 4 1/2 years (05-4318) Washington, DC: US Government Printing Office; 2006.

Odom, S. L. (2019). Peer-based interventions for children and youth with autism spectrum disorder: History and effects. *School Psychology Review, 48,* 170-176. https://doi.org/10.17105/SPR-2019-0019.V48-2

Odom, S. L., Cox, A., & Brock, M. (2013). Implementation science, professional development, and Autism Spectrum Disorders: National Professional Development Center on ASD. *Exceptional Children, 79*(3), 233-251. https://doi.org/10.1177/001440291307900207

Odom, S. L., Cox, A. W. , & Sam, A. M. (in press). *Autism Program Environment Rating Scale.* Baltimore, MD: Brookes Publishing.

Odom, S. L., Cox, A., Sideris, J., Hume, K. A., Hedges, S., Kucharczyk, S., Shaw, E., Boyd, B. A., Reszka, S., & Neitzel, J. (2018). Assessing quality of program environments for children and youth with autism: Autism Program Environment Rating Scale. *Journal of Autism and Developmental Disorders, 48*(3), 913-924. https://doi.10.1007/s10803-017-3379-7

Odom, S. L., Hume, K., Boyd, B., & Stabel, A. (2012). Moving beyond the Intensive Behavior Therapy vs. Eclectic dichotomy: Evidence-based and individualized program for students with autism. *Behavior Modification, 36,* 270-297. https://doi.org/10.1177/0145445512444595

O'nions, E., Happe, F., Evers, K., & Boonen, H. (2018). How do parents manage irritability, challenging behaviour, non- compliance and anxiety in children with autism spectrum disorders? A meta-synthesis. *Journal of Autism and Developmental Disorders, 48*(4), 1272-1286. https://doi.org/10.1007/s10803-017-3361-4

Pianta, R., La Paro, K. M., Hamre, B. K. (2008). *Classroom Assessment Scoring System: Pre-K version.* Baltimore, MS: Brookes Publishing Company.

Ruble, L., & McGrew, J. H. (2013). Teacher and child predictors of achieving IEP goals of children with autism. *Journal of Autism and Developmental Disabilities,* 43(12), 2748-2763. https//:doi. 10.1007/s10803-013-1884-x

Sam, A.M., Odom, S.L., Tomaszewski, B., Perkins, Y., & Cox, A. W. (2020). Employing evidence-based practices for children with autism in elementary schools. *Journal of Autism and Developmental Disorders,* (2020). Online prepublication. https://doi.org/10.1007/s10803-020-04706-x

Shih, W., Dean, M., Kretzmann, M., Locke, J., Senturk, D., Mandell, D. S., Smith, T. & Kasari C. (2019) Remaking recess intervention for improving peer interactions at school for children with autism spectrum disorder: Multisite randomized trial. *School Psychology Review, 48*(2), 133-144. https://doi:10.17105/SPR-2017-0113.V48-2

Simpson, R. , & Sasso, G. (1992). Full inclusion of students with autism in general education settings: Values vs. science. Focus on Autistic Behavior, 7(3), 1–13. https://doi.org/10.1177/108835769200700301

Soukakou, E. P. (2016). *Inclusive Classroom Profile.* Baltimore, MD: Brookes Publishing.

Tipton, E. (2014). How generalizable is your experiment? An index for comparing experimental samples and populations. *Journal of Educational and Behavioral Statistics, 39*(6), 478-501. https://doi.org/10.3102/1076998614558486

Tipton, E., Miller, K. (2016). *The Generalizer: A webtool for improving the generalizability of results from experiments.* Retrieved from http://www.thegeneralizer.org

U.S. Department of Education. (2019). *41st Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, 2019.* Washington, D.C.: Author.

Vivanti, G. (2020). Ask the Editor: What is the Most Appropriate Way to Talk About Individuals with a Diagnosis of Autism?. *Journal of Autism and Developmental Disorders, 50*(6), 691–693. https://doi.org/10.1007/s10803-019-04280-x

Wampold, B. E. (2015). How important are common factors in psychotherapy? An update. *World Psychiatry,14*(3). 270-277. https://doi.org/10.1002/wps.20238

Figure 1

Conceptual Framework for APERS

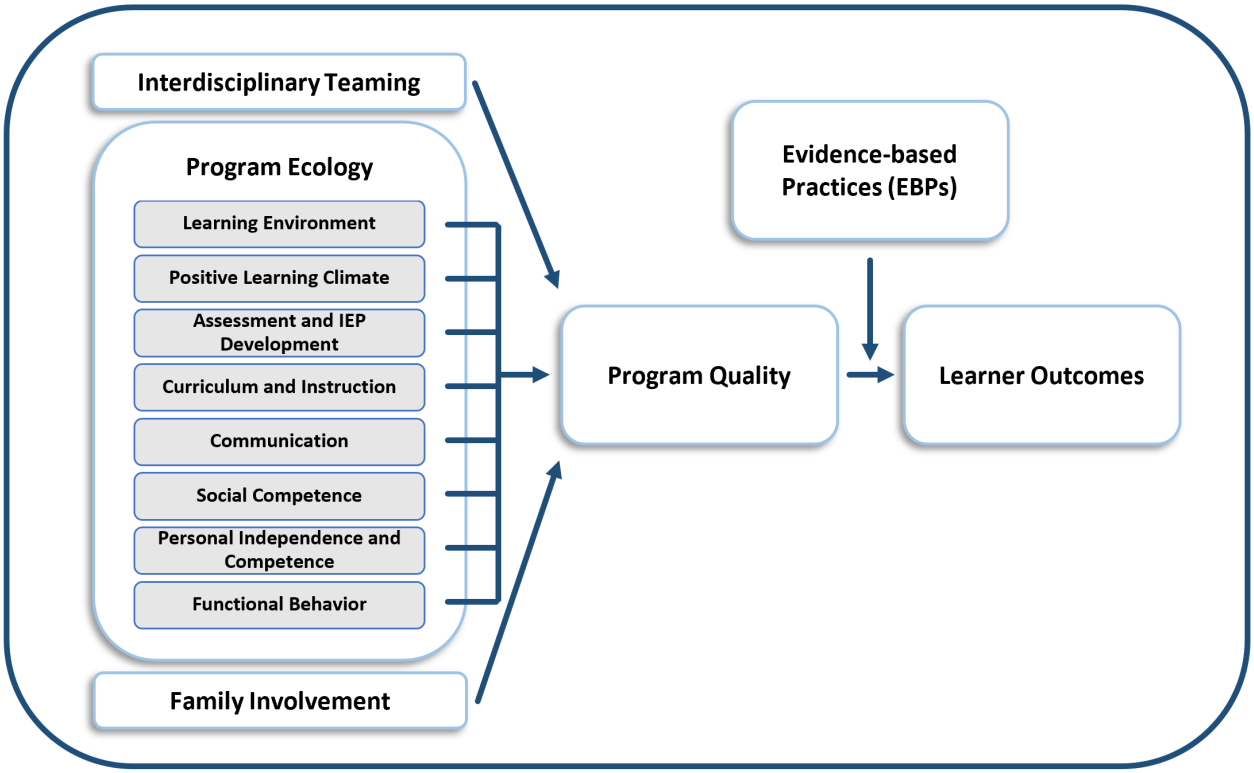

Figure 2

*Mean item rating for total score and domains with acceptable rating score (3.0) as referent. * indicates significantly different from 3.0 (minimally acceptable) rating.*

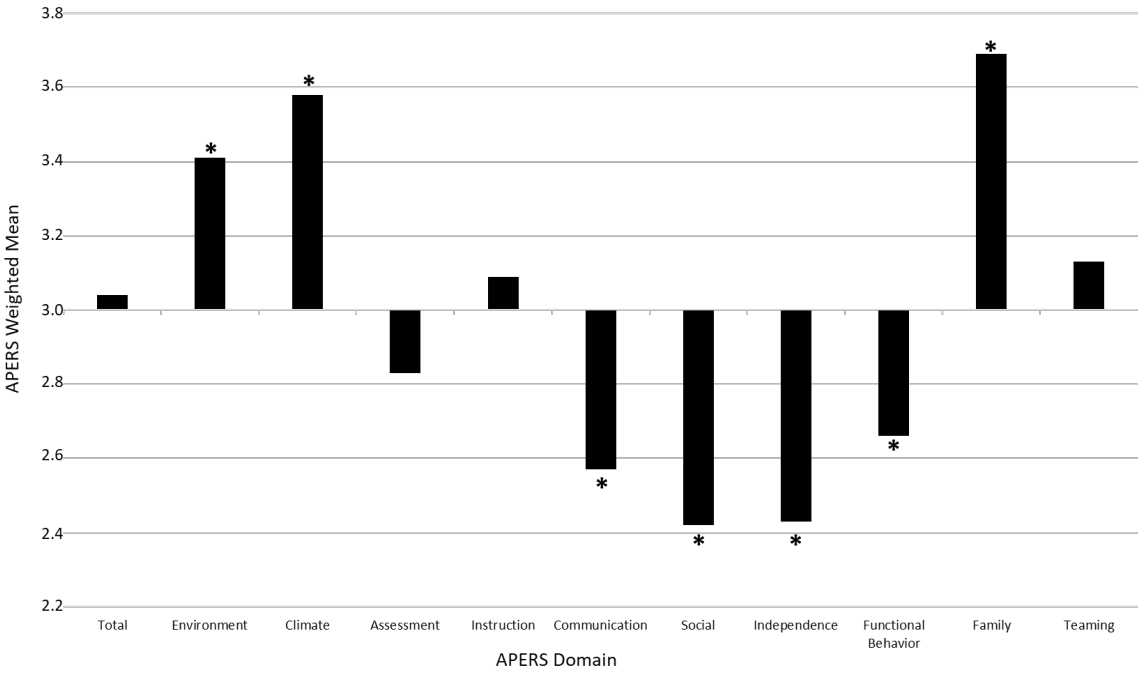

Figure 3

*APERS Total and Domain Ratings for IN and SE Programs. * indicates significant difference between groups.*

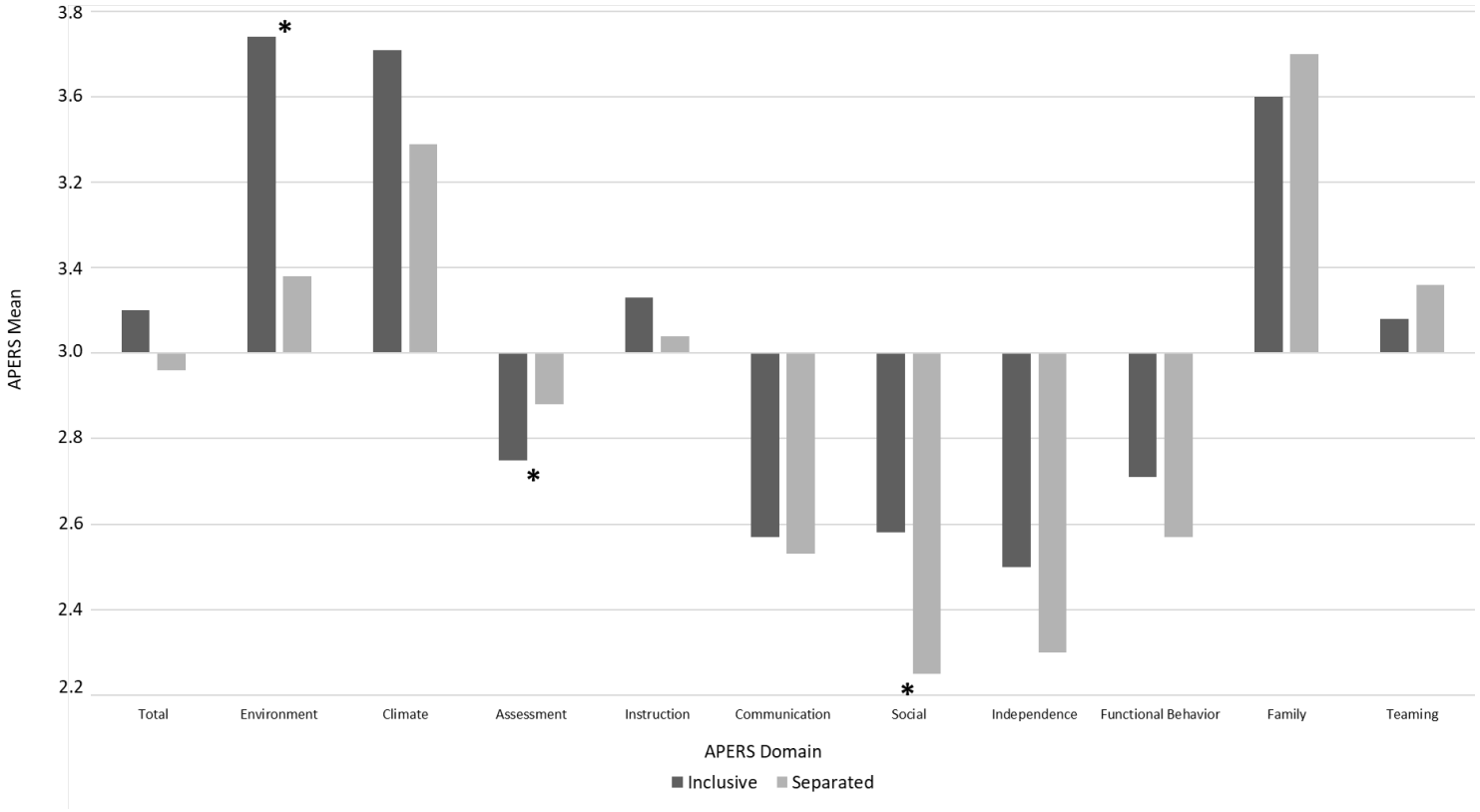

Table 1
Characteristics of Schools

Characteristics	% or M(SD)
Students receiving Free and Reduced Lunch	54.32
Race/Ethnicity	
White, Nonhispanic	44.86
Hispanic	18.72
Black, African American	26.15
Other	10.27
Title 1 Eligible	70.00
Average Number of Students	619.61(294.9)
Urbanicity	
City	44.50
Suburban	30.20
Rural	25.30

Table 2

Grade-level of Student Participants

Grade	Inclusive	Special Education	Total %
K	25	49	14.68
1	29	68	19.24
2	32	65	19.25
3	39	42	16.07
4	37	42	15.67
5	33	39	14.29
Other	0	4	0.79

Table 3

Cronbach Alphas and Intraclass Correlations for APERS-PE

	Standardized Cronbach Alpha			ICC	
	Weighted	Program Type		Program Type	
Domain		IN	SE	IN	SE
Learning Environment	0.85	0.78	0.85	0.93	0.98
Positive Learning Climate	0.79	0.82	0.68	0.54	0.94
Assessment and IEP Development	0.48	0.55	0.51	0.91	0.95
Curriculum and Instruction	0.89	0.88	0.86	0.87	0.95
Communication	0.75	0.59	0.77	0.71	0.80
Social Competence	0.72	0.67	0.71	0.93	0.93
Personal Independence and Competence	0.75	0.67	0.74	0.98	0.89
Functional Behavior	0.87	0.77	0.81	0.86	0.85
Family Involvement	0.76	0.69	0.79	0.91	0.93
Teaming	0.72	0.65	0.72	0.85	0.85
Overall	0.95	0.93	0.96	0.97	0.98

Appendix

APERS Preschool/Elementary Domain Content and Number of Items

APERS Domains	Description of Content	Number of Items
Learning Environments	Safety, organization, materials, visual schedules, transitions	9
Positive Learning Climate	Staff-student interactions, staff behaviors	4
Assessment and IEP Development	Assessing student progress, assessment process, IEP goals, transition planning	6
Curriculum and Instruction	Classroom instruction, focus on IEP goals, opportunity to generalize, prompting, accommodations	12
Communication	Planning for communication, communication rich environment, individualized communication instruction, responsiveness to student, communication systems.	4
Social Competence	Arranging opportunities, teaching and modeling, personal hygiene and relationships, social skills training, peer social networks,	4
Personal Independence	Self-advocate for accommodations, self-management, choices available	4
Functional Behavior	Proactive strategies, behavioral assessment, data collection, teaming	5
Family Involvement	Teaming, communication, parent teacher meetings,	4
Teaming	Team membership, team meetings, decision making,	7