June 2020 Education and Training in Autism and Developmental Disabilities Vol. 55, No. 2, pp. 117–244

Education and Training in Autism and Developmental Disabilities

Focusing on individuals with
autism, intellectual disability and other developmental disabilities

DADD

Volume 55 Number 2 June 2020

Education and Training in Autism and Developmental Disabilities

The Journal of the Division on Autism and Developmental Disabilities, The Council for Exceptional Children

Education and Training in Autism and Developmental Disabilities is sent to all members of the Division on Autism and Developmental Disabilities of The Council for Exceptional Children. All Division members must first be members of The Council for Exceptional Children. Division membership dues are $30.00 for regular members and $15.00 for full time students. Membership is on a yearly basis. All inquiries concerning membership, subscription, advertising, etc. should be sent to the Division on Autism and Developmental Disabilities, 3100 Clarendon Blvd., Suite 600, Arlington, Virginia 22201-5332. Advertising rates are available upon request.

Manuscripts should be typed, double spaced, and emailed (etadd@asu.edu) to the Editor: Stanley H. Zucker, Mary Lou Fulton Teachers College, Box 871811, Arizona State University, Tempe, AZ 85287-1811. Each manuscript should have a transmittal letter and a cover sheet that gives the names, affiliations, and complete addresses of all authors.

Editing policies are based on the **Publication Manual,** the American Psychological Association, 2009 revision. **Additional information is provided on the inside back cover.** Any signed article is the personal expression of the author; likewise, any advertisement is the responsibility of the advertiser. Neither necessarily carries Division endorsement unless specifically set forth by adopted resolution.

Education and Training in Autism and Developmental Disabilities is abstracted and indexed in **Psychological Abstracts, PsycINFO, e-psyche, Abstracts for Social Workers, International Journal of Rehabilitation Research, Current Contents/Social and Behavioral Sciences, Excerpta Medica, ISI Social Sciences Citation Index, Adolescent Mental Health Abstracts, Educational Administration Abstracts, Educational Research Abstracts,** and **Language and Language Behavior Abstracts.** Additionally, it is annotated and indexed by the ERIC Clearinghouse on Handicapped and Gifted Children for publication in the monthly print index **Current Index to Journals in Education** and the quarterly index, **Exceptional Child Education Resources.** Access is also available in EBSCO, ProQuest, and JSTOR.

Education and Training in Autism and Developmental Disabilities Vol. 55, No. 2, June 2020, Copyright 2020 by the Division on Austim and Developmental Disabilities, The Council for Exceptional Children.

Division on Autism and Developmental Disabilities

The Council for Exceptional Children

The purposes of this organization shall be to advance the education and welfare of persons with autism and developmental disabilities, research in the education of persons with autism and developmental disabilities, competency of educators in this field, public understanding of autism and developmental disabilities, and legislation needed to help accomplish these goals. The Division shall encourage and promote professional growth, research, and the dissemination and utilization of research findings.

EDUCATION AND TRAINING IN AUTISM AND DEVELOPMENTAL DISABILITIES (ISSN 2154-1647) (USPS 0016-8500) is published quarterly, by The Council for Exceptional Children, Division on Autism and Developmental Disabilities, 3100 Clarendon Blvd., Suite 600, Arlington, Virginia 22201-5332. Members' dues to The Council for Exceptional Children Division on Developmental Disabilities include $8.00 for subscription to EDUCATION AND TRAINING IN AUTISM AND DEVELOPMENTAL DISABILITIES. Subscription to EDUCATION AND TRAINING IN AUTISM AND DEVELOPMENTAL DISABILITIES is available without membership; Individual—U.S. $100.00 per year; Canada, PUAS, and all other countries $104.00; Institutions—U.S. $249.00 per year; Canada, PUAS, and all other countries $254.00; single copy price is $40.00. U.S. Periodicals postage is paid at Arlington, Virginia 22204 and additional mailing offices.

POSTMASTERS: Send address changes to EDUCATION AND TRAINING IN AUTISM AND DEVELOPMENTAL DISABILITIES, 3100 Clarendon Blvd., Suite 600, Arlington, Virginia 22201-5332.

Education and Training in Autism and Developmental Disabilities

VOLUME 55 NUMBER 2 JUNE 2020

Manuscripts Accepted for Future Publication in Education and Training in Autism and Developmental Disabilities

September 2020

Paraprofessional-implemented systematic instruction for students with disabilities: A systematic literature review. **Virginia L. Walker**, Karen H. Douglas, Sarah N. Douglas, and Sophia R. D'Agostino, UNC Charlotte, Dept. of Special Education & Child Development, 9201 University City Blvd., Charlotte, NC 28223.

Secondary transition practices and predictors for students with autism spectrum disorder. David W. Test, Erik Carter, **Jennifer Coyle**, Rachel Seaman-Tullis, Dana Rusher, and Samuel L. Odom, Western Michigan University, 1903 W. Michigan Ave., Sangren Hall, Kalamazoo, MI 49008.

Evidence for a literacy intervention for students with intellectual and developmental disabilities. **Jill H. Allor**, Paul Yovanoff, Stephanie Al Otaiba, Miriam B. Ortiz, and Carlin Conner, Southern Methodist University, Department of Teaching and Learning, Simmons School of Education and Human Development, Post Office Box 750455, Dallas, TX 75275-0455.

Dissemination and decision-making: Factors related to pre-service practitioners' selection of practices for students with autism. Maria L. Hugh, **LeAnne Johnson**, and Veronica P. Fleury, Educational Psychology, 56 East River Pkwy, Education Sciences Building 243, University of Minnesota, Minnesota, MN 55455.

Increasing the provision of choices within an adult transition program. **Rachel J. Schwartz**, Rachel E. Robertson, and Sarah Westerfield, Department of Instruction and Learning, University of Pittsburgh, 5529 Wesley W. Posvar Hall, 230 South Bouquet Street, Pittsburgh, PA 15260.

Effects of general education teacher-delivered embedded instruction in teaching students with intellectual disability to solve word problems. **Jessica Bowman**, John McDonnell, Joanna Ryan, Olivia Fudge Coleman, Lyndsey Aiono Conradi, and Carrie Eichelberger, University of Utah, 1721 Campus Center Drive, SAEC 2280, Salt Lake City, UT 84112.

Enhancing early numeracy skills of children with severe disabilities and complex communication needs. **John C. Wright**, Victoria F. Knight, Christopher J. Lemons, Esther Lindström, and Julia Strauss, Peabody College of Vanderbilt University, Department of Special Education, 110 Magnolia Circle, Nashville, TN 37203-5721.

What is the proof now? An updated methodological review of research on social stories. Christine M. Milne, **Justin B. Leaf**, Joseph H. Cihon, Julia L. Ferguson, John McEachin, and Ronald Leaf, Autism Partnership Foundation, 200 Marina Drive, Seal Beach, CA 90740.

Address is supplied for author in **boldface** type.

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 119–141

Studies Comparing Augmentative and Alternative Communication Systems (AAC) Applications for Individuals with Autism Spectrum Disorder: A Systematic Review and Meta-Analysis

Orhan Aydin and Ibrahim H. Diken
Anadolu University

Abstract: In the present study, we aimed to make an evaluation, conduct comprehensive descriptive analyses, and calculate the effect sizes of studies which compare AAC applications designed according to single-subject research designs and conducted for individuals with ASD. Within the scope of this aim, 21 studies meeting the inclusion criteria were determined among the studies obtained by the systematic review. Firstly, the descriptive analysis of the included studies was conducted, and then they were evaluated in terms of the "Quality Indicators of Single-Subject Experimental Research Studies" and "acceptability" criteria suggested by Horner et al. (2005). The improvement rate difference effect sizes of 11 studies that were assessed as "acceptable" were calculated, and they were evaluated according to the design standards of the What Works Clearinghouse (WWC, 2017). The obtained findings were discussed in the light of the literature, and various suggestions were presented for future studies and practices.

Autism spectrum disorder is a neuro-developmental disorder of which estimated prevalence has increased significantly in the last 20 years (Rice et al., 2012). The prevalence, which was 4–5 per 10000 in the early 1990s, is stated to be 1 per 68 according to the data for 2010 (Centers for Disease Control and Prevention-CDC, 2014), and 1 per 59 according to the data for 2014 (Baio et al., 2018). Nowadays, the increasing number of individuals diagnosed with ASD has led to an increase in the number of researchers working on the education of these individuals. Researchers conduct studies on various skills such as communication skills of these children, academic skills, independent living and reducing problem behaviors (e.g. Boyd et al., 2015; Caroll & Kodak, 2014; Fletcher-Watson et al., 2016; O'Malley et al., 2014; Schmidt et al., 2014).

The majority of individuals with ASD draw attention with their deficiencies in communication skills, which are an important criterion in the diagnostic criteria included in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; APA, 2013). In the literature, it is stated that approximately 50% of these individuals cannot speak fluently (Hart & Banda, 2009). Furthermore, among individuals with ASD who attend a program in an educational environment, between 25% and 61% have very little communication skills or do not have functional communication skills (Schlosser & Wendt, 2008). However, in the literature, there are different applications to increase communication skills of these individuals. Among these applications, augmentative and alternative communication systems (AAC) are used by many researchers and practitioners to increase communication skills of individuals with ASD (Ganz et al., 2012). AAC applications are temporary or continuous practices aimed at eliminating or supporting deficiencies in the language, speech, and communication skills of individuals (Vento-Wilson, 2014). AAC applications are generally classified in two ways as aided and unaided. While unaided applications are the applications in which an individual communicates by

Correspondence concerning this article should be addressed to Orhan AYDIN, Research Institute for Individuals with Disabilities, Anadolu University, Eskisehir, 26210, TURKEY. E-mail: o_aydin@anadolu.edu.tr

using his/her own body (e.g. manual sign, gestures, and mimics), aided applications are the applications in which an external intervention (e.g. images, words, speech-generating technological devices) is performed (Mirenda, 2003; Mueller, 2014).

Unaided AAC applications, especially manual sign, are the applications that can be used for communication. However, the fact that the listener does not know a process like a manual sign and there are not individuals who can understand and respond to the manual sign in every environment (e.g. store, restaurant) is a limitation regarding its functionality. Furthermore, the inadequacy of motor skills and the limited ability to use gestures make it challenging to learn unaided AAC applications for individuals with developmental disabilities such as ASD (Lorah et al., 2015; Mirenda, 2003). Aided AAC applications include a variety of applications such as Picture Exchange (PE), Picture Exchange Communication Systems (PECS), Speech Generating Devices (SGD) or Voice Output Communication Aids (VOCA), high-tech speech generating device applications running on iPad or tablet (e.g. Proloquo2go, MyTalk; Mueller, 2014; Shane et al., 2012).

Aided AAC applications are practices that can be used to increase some academic skills of individuals with ASD, especially communication skills, and to reduce their behavioral problems. Among these applications, PECS, which include picture-based communication processes, were developed by Bondy and Frost (1994). In this application, the child exchanges the desired object with the picture by giving the picture of the object that he/she desires to the person with whom he/she communicates. The PECS is an advantageous application because it requires fewer motor skills. In a study conducted by Tincani (2004), it was found out that the PECS was more effective than the manual sign practice in communication skills of a child who had inadequate motor skills. Furthermore, PECS practices are less costly than SGD applications. Besides these advantages, it is considered a disadvantage that these practices do not include an educational process related to speech output (Boesch et al., 2013a). SGD applications are portable technological devices that contain symbols that can express a word or sentence and that can provide digital or artificial sound output. SGD applications are durable, in contrast to PECS in which the used materials are damaged over time. They are practices that are simple to use and focus more on sound output (Boesch et al., 2013a). Because of these characteristics, it is stated that the SGD aplications are more preferred by individuals with ASD (e.g. Achmadi et al., 2014; Couper et al., 2014; Lorah, 2016). Besides these benefits, they are more expensive (Boesch et al., 2013a, Lorah et al., 2015), and they are heavy and bulky in terms of portability, especially for individuals with ASD who have a physical disability (Boesch et al., 2013a).

A large number of studies have been carried out in the literature on aided and unaided AAC applications. In addition, review and meta-analysis studies have been conducted in order to evaluate what kind of application is more effective in teaching what kind of skills in individuals with ASD. For example, in the review and meta-analysis study conducted by Schlosser and Wendt (2008), 11 studies, nine of which were single-subject and two of which were experimental, were evaluated. While six of the studies were PECS practices, three of them were on SGD. Furthermore, both PECS and manual sign were applied in one study. One study includes only unaided AAC applications. As a result of their analysis, the researchers concluded that AAC applications did not prevent individuals with ASD from generating speech, on the contrary, these applications supported speech generation.

In the meta-analysis study conducted by Ganz et al. (2012), 24 studies designed according to the single-subject research design, in which aided AAC applications were applied, were reviewed. While 16 of these studies were picture-based or PECS applications, eight of them were applications related to SGD. According to the results of the Improvement Rate Difference (IRD) effect size (84% CI), it is stated that aided AAC applications are effective applications on the target behavioral outcomes of individuals with ASD. Moreover, although aided AAC applications also affect other skills (e.g. academic skills, reduction of problem behaviors), it has been reported that

they have an effect especially on communication skills of individuals with ASD.

In a systematic review study conducted by Logan et al. (2017), 30 studies consisting of 24 single-subject and six group experimental studies, in which aided AAC applications were performed to increase the social communication skills of individuals with ASD, were evaluated. Furthermore, the researchers examined applications in terms of immediate effect, and also generalization, maintenance, and social validity. Of the studies carried out, 17 were studies based on the picture (photograph) or PECS. 12 of the studies were applications with SGD in which devices such as iPad and iPod could also be used. One of the studies was the JASPER application developed and implemented by Kasari et al. (2014). Although the findings demonstrated that aided AAC applications increase the social interaction skills of individuals with ASD, it was concluded that in most of the studies generalization, maintenance, and social validity data were not considered. Therefore, it was suggested to focus on the long-term effects and generalizability of aided AAC applications in future studies.

In the another meta-analysis study conducted by Muharib and Alzrayer (2017), high-tech SGD applications aimed at individuals with ASD were analyzed. 20 single-subject studies (95% confidence interval), carried out by using tools and applications such as GoTalk, Proloquo2Go, iPad, iPod as high-tech applications, were analyzed by calculating the IRD. The findings demonstrated that high-tech applications are practices that have a significant impact on the teaching of the skills of requesting, intraverbal and multistep tacting to children with ASD. At the same time, the studies were evaluated according to the quality indicators proposed by Horner et al. (2005), and it was stated that high technology SGD applications had a moderate effect.

There is no specific meta-analysis and systematic review of the studies conducted for the purpose of reviewing comparative AAC applications. In the literature, there is a need for a systematic review of these studies because of the presence of a significant number of these comparative applications and the fact that most of them are conducted according to single-subject research designs. In the present research, we aimed to review studies of comparative AAC applications with single subject research designs for individuals with ASD. Therefore, in the light of the obtained findings, it was aimed to determine which AAC applications could be used effectively in the teaching of social communication or other skills. Furthermore, we aimed to evaluate studies in terms of quality indicators and design standards and calculate effect sizes of them. Within the scope of these purposes, we aimed to (a) perform descriptive analyses of all studies included, (b) evaluate them in terms of the "Quality Indicators of Single-Subject Experimental Research Studies" and "acceptability" criteria suggested by Horner et al. (2005), (c) calculate to effect size of studies which were assessed as "acceptable" and evaulate according to What Works Clearinghouse (WWC, 2017) design standards of them. In this study, it was preferred to use the quality indicators proposed by Horner et al. since these indicators are more comprehensive than other rubrics (e.g. treatment fidelity, participant characteristics are not a criterion in the WWC). However, Horner et al. do not consider single-subject research designs in terms of graphical quality. On the contrary, graphical quality is examined in the WWC design standards. Therefore, the relevant part in the WWC was used for investigating graphical quality in this study.

Method

Search Procedure

To access studies, Academic Search Complete, ERIC, MEDLINE, PsycARTICLE, Science Citation Index, ScienceDirect, and Scopus databases were reviewed. We involved in studies which were carried out with individuals with ASD and which were published in the English language in international peer-reviewed journals between January 1980 and November 2017. The review was started from 1980, since, in a separate diagnostic category, ASD appeared in the DSM-3 (APA, 1980), which was first published in 1980. In the process of reviewing the electronic databases, the terms *"autism, Asperger Syndrome, Autism Spectrum Disorder, PDD-NOS"* were used in relation to the term of ASD. In relation to AAC applications, the words *"functional communication, picture ex-*

change communication system, PECS, picture exchange, manual sign, augmentative communication, alternative communication, communication systems, communication aids, augmentative and alternative communication, AAC, speech generating device, voice output communication aid" were used. As a result of the review, 177 studies that were considered to be related to AAC applications were downloaded to the computer and filed in electronic format. Moreover, two additional studies were also filed in the computer by examining the references of the studies which met the comparative study criteria. Moreover, the citation lists of the studies which met the comparative study criteria were accessed by using the Google Scholar search engine, the headings of the studies on these lists were examined, and their abstract sections were reviewed if required, and ultimately two additional studies were filed by being downloaded. Thus, a total of 181 studies were filed.

Inclusion/Exclusion Criteria

In this analysis study, the inclusion criteria were determined as follows: (a) being published in the English language in international peer-reviewed journals between 1980 (January) and 2017 (November), (b) having at least one of the participants diagnosed with ASD (autism, Asperger syndrome, pervasive developmental disorder), (c) comparing AAC applications, and (d) being designed according to single-subject research designs. The exclusion criteria were determined as follows: (a) using a research design other than single-subject research designs, (b) studying different disability groups other than individuals with ASD, (c) being literature reviews, (d) not comparing AAC applications, and (e) being studies that do not provide appropriate data for visual analysis. Among the obtained 177 studies, group experimental studies (n = 19), unrelated studies not including the AAC application process (n = 18), case studies (n = 14), systematic review and meta-analysis studies related to ASD and other forms of disability (n = 12), master's theses or doctoral dissertations (n = 6), AAC applications conducted for individuals without ASD (n = 4), quantitative studies (n = 4), design-based studies (n = 3), informative studies on AAC applications (n = 3), mixed design studies (n = 2), descriptive studies (n = 2), longitudinal study (n = 1) and predictive study (n = 1) were excluded from the research. Among the obtained studies, single-subject research designs (n = 88) were reviewed again, and it was determined that there were a total of 22 comparative studies. The reference lists and Google Scholar citation lists of comparative studies were reviewed, and four studies were added. Therefore, the sum of comparative studies became 26. One of these studies (Hill & Flores, 2014) was eliminated since it did not provide appropriate data for visual analysis, and another one (Soto et al., 1993) was eliminated since it was carried out for individuals without ASD. Moreover, two studies (Scholosser et al., 2007; Trembath et al., 2009) were eliminated since they were comparative studies using other methods/applications instead of the AAC application, and another study (Ganz et al., 2014) was eliminated because it did not contain an educational process related to AAC applications and aimed to determine the frequency of use of two different AAC applications. The remaining 21 studies were filed to be included in the advanced analysis processes, by the consensus of the researchers. The flow of the processes within the scope of the search procedure and the inclusion/exclusion criteria are presented in Figure 1.

Analysis Process

All of the included studies were first analyzed descriptively, and they were evaluated in terms of the *"Quality Indicators of Single-Subject Experimental Research Studies"* suggested by Horner et al. in 2005. Then, concerning the quality indicators, the *"acceptable"* studies were determined. The meta-analyses of the studies assessed as *"acceptable"* were conducted, and they were also evaluated regarding the WWC (2017) design standards.

Descriptive Analysis. In the descriptive analysis process, the studies were coded in terms of the variables of the number of participants, age, gender, diagnosis, and intelligence quotient (IQ). The studies were coded in terms of the variables of the setting in which studies were conducted, the instructional arrangement, the design of the research, the depen-

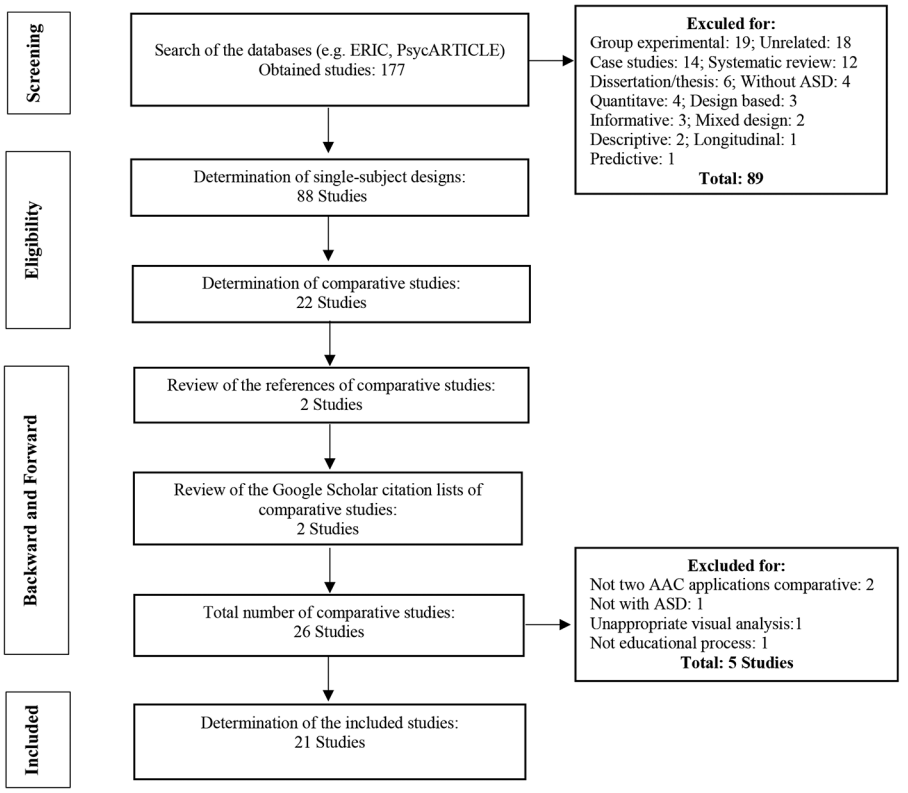


Figure 1. The evaluation of studies within the scope of the search process and inclusion/exclusion criteria.

dent variable, the compared AAC applications, the other contents used in the application process (e.g. least to most prompting, discrete trials teaching, etc.), the preferred application, the practitioner information, generalization, maintenance and social validity, interobserver agreement for the dependent variable, and treatment fidelity with regard to the practice process.

Quality Indicators. In a single-subject research, Horner et al. (2005) mentioned a total of 21 quality indicators in five categories as to be (a) description of participants and the setting, (b) dependent variable, (c) independent variable, (d) baseline, and (e) validity (internal, external and social validity). Since a study is evaulated as an *"acceptable"* study in terms of the *"Quality Indicators of Single-Subject Experimental Research Studies"* developed by Horner et al. (2005), information should be given about the listed five characteristics in that study: (a) operational definition of the application, (b) functional definition of findings, (c) treatment fidelity, (d) functional relationship between the application and the findings obtained, and (e) carrying out the experimental control with a sufficient number of studies, researchers, and participants. There is no consensus in the literature on which one of these quality indicators that are coded according to these assessment criteria will be considered. In this study, six quality indicators (see; Table 2 items) were taken into consideration including: for *"(a) the operational definition of the application"* among quality indicators item 9, for *"(b) the functional definition of findings"* items 14 and 15, for *"(c) treatment fidelity"* item 11, for *"(d) the functional relationship between the application and the findings obtained"* item 16, and for *"(e) carrying out the experimental control with a sufficient number of studies, researchers, and participants"* item 17. Which items to take into consideration from the quality indicators (Table 2) regarding the first four characteristics (a, b, c and d) we decided in accordance with the recommendations of Horner et al. as in

the work of Aydın et al. (2019). However, with regard to the statement *"carrying out the experimental control with a sufficient number of studies, researchers, and participants"*, Horner et al. (2005) mentioned that *"at least five studies should be carried out with regard to the application, studies should be conducted in at least three different regions by three different research groups, and they should be carried out with at least 20 participants."* As the studies were not based on a single independent variable, we accepted to evaluate them in accordance with item 17 by deciding that it was appropriate to evaluate this feature in itself of the studies (see, Table 2; *: Indicators that are considered as the "acceptability" criteria).

WWC Design Standards. The WWC (2017) seeks the following criteria in the design standards of single-subject research: (a) For reversal/withdrawal (AB) designs, at least four phases must be designed per case, and there must be at least five data points per phase to be rated *"meet the WWC design standards."* In case of presenting at least three data in each phase, it is rated *"meet design standards with reservations."* In case of presenting less than three data, graphical data do not meet the WWC design standards. (b) For multiple baseline and multiple probe designs, there must be at least six phases in total and at least five data points in each phase to be rated *"meet the WWC design standards."* In case of the presentation of at least three data in each phase, it is rated *"meet design standards with reservations"*. Graphical data do not meet the WWC design standards if less than three data are presented in one of the phases. (c) For alternating treatment designs, there must be at least five times (e.g. BCBCBCBCBC) comparative data for per condition in each phase and there should be a comparison of maximum two applications in each phase to be rated *"meet the WWC design standards"*. In case of the presentation of comparative data four times for per condition, it is rated *"meet design standards with reservations."* In the case of a comparison of more than two applications, a comparison of each application with each other should be made separately (e.g. comparison of A and B, comparison of A and C, comparison of C and B). If there is a comparison of more than two conditions in the same phase, it does not meet the WWC design standards. In this study, the comparative studies (mostly alternating treatment designs) that are considered as *"acceptable"* by Horner et al. (2005) were evaluated in terms of the WWC design standards.

Data Extraction. We saved screenshots of the graphs in the studies on the computer for the effect size calculation. Then, we transferred them to the GraphClick software program to digitize data. GraphClick is a software program with high accuracy and reliability in digitizing graphical data (Boyle et al., 2013; Rakap et al., 2016). We transferred graphical data to Windows Excel after digitizing with the GraphCilck software program, and then, we calculated the effect sizes by entering the digitized data to the related columns from the address of www.singlecaseresearch.org for the calculation of effect sizes.

Effect Size Calculation. In the study, in order to determine and compare the effectiveness of AAC applications, the effect size was calculated using the improvement rate difference (IRD) technique (Parker et al., 2009). The IRD is an analysis method adapted from the medical field and defined as the difference in the amount of progress between phases A and B (Parker et al., 2009). The IRD has many advantages such as: (a) interpreting the difference in the rate of improvement between the baseline and intervention phases appropriately, (b) being performed with a simple hand calculation, (c) being consistent with the visual analysis, (d) achieving the confidence intervals easily, (e) not seeking for prerequisites in distribution assumptions (e.g. normal distribution), and (f) being implemented to analyze complex single-subject research designs and multiple data series (Parker et al., 2009). Furthermore, considering confidence intervals is a unique guiding approach in the effect size calculation of single-subject research designs (Kratochwill et al., 2010). Besides these advantages, the IRD has a more sensitive measure in determining differences between the groups (Ganz et al., 2012). Due to this usefulness, the effect size was calculated using the IRD analysis technique in this study.

We calculated IRD scores using web-based IRD calculator at www.singlecaseresearch.org (Vannest et al., 2016). The obtained IRD values vary between 0 and 1. If the value of IRD is 0.50 or less, it is interpreted as a small or

suspect effect, if it is between 0.50–0.70, it is interpreted as a medium effect, and if it is higher than 0.70, it is interpreted as a high effect. In addition to the IRD, calculations in the 95% confidence interval were also performed to determine the level of precision of the IRD values. The confidence interval was 95% in this study because using the confidence interval of 84% was considered as a limitation in the study by Ganz et al. (2012), and they have suggested using high confidence interval levels in future meta-analysis studies. While confidence intervals close to each other mean that the calculated IRD values are more reliable and valid, wide confidence intervals mean that the validity of the IRD values is low (Parker et al., 2009). For example, when the calculated IRD value in the 95% confidence interval is (.80, .95) .90, it means that the IRD values are between the values of .80 and .95 with a probability of 95%, in other words, close ranged. When the calculated IRD value in the 95% confidence interval is (0.10, 0.62) .35, it means that the IRD values are between the values of 0.10 and 0.62 with a probability of 95%, in other words, wide ranged.

Reliability

We obtained five types of reliability data being related to, (a) descriptive analysis, (b) *"Quality Indicators of Single-Subject Experimental Research Studies"*, (c) data extraction with GraphClick, (d) effect size of the improvement rate difference, and (e) confidence interval calculation processes. In this context, in the first stage, 33.3% of the studies ($n = 7$) were determined to be random, and for the first two reliability analyses (a and b), these studies were used. Then, 36.3% of the studies evaluated as *"acceptable"* ($n = 4$) were determined to be random, and for the last three reliability analyses (c, d, and e), these studies were used. The reliability analysis was carried out by an independent researcher, who is experienced in the mentioned processes, over the studies which were determined to be random. In the coding for these mentioned processes, consistency between coders was calculated. Therefore, the reliability data between coders were obtained by dividing the number of variables (such as indicators, numerical values of data points, evaluation points) in which an agreement was provided by the total number of variables with agreement and disagreement, and then by multiplying the result by 100. Reliability data between coders were obtained to be 94% (range: 87%–100%) for descriptive analysis, 96% (86%–100%) for quality indicators, and 100% for data extraction with the GraphClick software program, improvement rate difference, and confidence interval calculations. The coders came together and discussed the variables they could not match and reached a common decision.

Results

Results of Descriptive Analysis

The coding results of the studies ($n = 21$) regarding the related variables (e.g. dependent variable, research design) are presented in Table 1. The studies were carried out with a total of 69 children/adolescents with ASD (individuals with autism, pervasive developmental disorder or childhood disintegrative disorder). In 20 of the studies, information was provided about the gender of 60 participants, and 49 of them were male, and 11 of them were female. In the studies, which presented information about the age of participants ($n = 20$), while there were 31 participants aged between 3–6 years (pre-school period), 26 participants were between the ages of 6.1–11 years (primary school period), and 9 participants were between the ages of 11.1–14 years (secondary school period). Furthermore, in a study (Beck et al., 2008), the ages of participants were not presented; instead, it was reported that all participants ($n = 3$) were in the pre-school period. Since two separate studies were published from a doctoral dissertation study and three participants in those studies (Boesch et al., 2013a, b) were the same persons, they were not counted two times.

In terms of setting in which the studies were conducted, it is seen that more than one setting could be used in a study. While in some of these studies ($n = 11$; e.g. Tincani, 2004), practices were conducted in a quiet corner of the classroom in the special education class/classroom environment, some of them ($n = 9$; e.g. Achmadi et al., 2014) were conducted in the home environment (e.g. kitchen, dining

TABLE 1

Descriptive Analysis of the Studies

Study	*Participant (Number, age, gender, diagnosis)*	*Setting/Instructional Arrangement*	*Research Design*	*Dependent Variable*	*Compared AAC Applications*	*Instructional Contents*	*Preferred Application*	*Practitioner*	*G/M/SV*	*IOA/TF*
Achmadi et al. (2014)	3/4-5-5/3 M/Autism(3)	Home environment(2), classroom environment(1)/1:1	ATD	Request	MS, PE, SGD	Discrete trial teaching, graduated guidance, progressive time delay procedure	SGD(3)	Researcher	N/Y/N	Y/Y
Agius & Vance (2016)	3/3.1-4.3-4.5/3 M/ Autism (3)	Intervention room/two practitioners together	MBD with ATD	Request	PECS, SGD with iPad	PECS protocol, error correction	SGD(3)	Researcher	N/N/Y	Y/Y
Beck et al. (2008)	3/no age all three pre-school/3 M/ Autism (2), PDD-NOS(1)	Special education class/two practitioners together	ATD	Request, vocalization/ verbalization, using PECS and VOCA	VOCA, PECS	PECS protocol	–	Researcher	Y/N/N	Y/Y
Boesch et al. (2013a)*	3/6-7-10/2 M,1 F/Autism (3)	Therapy room(2), Kitchen at home(1)/ three practitioners together	MBD with ATD	Request	SGD, PECS	PECS protocol, using prompt	–	–	N/Y/Y	Y/Y
Boesch et al. (2013b)*	3/6-7-10/2 M, 1 F/Autism (3)	Clinic(2), home environment(1)/–	MBD with ATD	Social-com., natural speech production	SGD, PECS	PECS protocol,	–	–	N/Y/N	Y/Y
Chen et al. (2016)	3/12–13-13/3 M/ ASD (3)/IQ: 53–55-62	Teaching room/–	MT-RD (ABCACB)	Request, greeting, accurately responded to questions	SGD with two different interfaces	–	–	–	N/N/N	N/N
Couper et al. (2014)	9/between 4.2–12.3/ASD(9)	In a quiet room at school(3), in the kitchen or living room at home(5), clinic(1)/1:1	ATD	Request	MS, PE, SGD	Discrete trial teaching, 10 s. time delay procedure, graduated guidance	SGD(8), None of them(1)	Therapist, child psychologist, Ph.D. student, special education teacher	N/Y/N	Y/Y
Gevarter et al. (2017)	5/3.1-4.4-4.6-6.3-8.8/3M-2F/ASD(5)	Home environment./–	ATD	Request	SGD with four different picture contents	6 s. time delay procedure, gradually increasing the prompt	–	Graduate students	N/N/N	Y/Y
Gevarter et al. (2014)	3/3.1-3.6-3.11/3M/ ASD(3)	Home environment/–	ATD	Mands	SGD with three different picture contents with iPad	6 s. time delay procedure, The least to most prompting procedure	–	Researcher, special education teacher	N/N/N	Y/Y
Lorah (2016)	5/8.11–8.5-9.9-10.5-12.7/4 M-1 F/ Autism (5)	Special education class/1:1	ATD	Mands	PE, SGD	5 s. time delay procedure, error correction	SGD(4)	Teacher, teacher assistant	N/N/Y	Y/Y

TABLE 1—(*Continued*)

Study	*Participant (Number, age, gender, diagnosis)*	*Setting/Instructional Arrangement*	*Research Design*	*Dependent Variable*	*Compared AAC Applications*	*Instructional Contents*	*Preferred Application*	*Practitioner*	*G/M/SV*	*IOA/TF*
Lorah et al. (2013)	5/3.10-4.1-4.3-5.5-5.11/5 M/Autism (5)	Special education class/–	ATD	Mands	PE, SGD with iPad	5 s. time delay procedure	SGD(4), PE(1)	Two master students One Ph.D. student	N/Y/N	Y/Y
McLay et al. (2015)	4/5.2-7-8-10.1/3 M-1 F/Autism (4)	Special education class or adjoining room/1:1	ATD	Request	MS, PE, SGD	10 s. time delay procedure	SGD(4)	Teacher assistants, speech-language pathologist, researcher	Y/Y/N	Y/Y
McLay et al. (2017)	2/5.4-10.3/2 M/ASD(2)	Therapy room/1:1	MPD with ATD	Request	MS, PE, SGD	10 s. time delay procedure, graduated guidance	SGD(2)	–	N/N/N	Y/Y
Schlosser & Blischak (2004)	4/8-9-12-12/4 M/Autism (4)	A quiet corner of the classroom/two practitioners together	ATD (adapted)	Spelling	Print, speech and print-speech SGD	Copy-Cover-Compare	SGD_{print}(4)	Two research assistants	Y/Y/Y	Y/Y
Schlosser et al. (1998)	1/10 years/M/Autism/IQ=80	School library, a quiet corner of the classroom/two practitioners together	ATD (adapted)	Spelling	Visual, auditory and visual-auditory SGD	Copy-Cover-Compare	$SGD_{audotiry}$	Teacher and teacher assistant	N/Y/Y	Y/Y
Son et al. (2006)	3/3-3.8-5.5/2 F, 1 M/Autism (2), PDD-NOS (1)	The kitchen at home/1:1	ATD	Request	VOCA, PE	The least to most prompting procedure	PE(2), VOCA(1)	Researcher	N/N/N	Y/N
Tincani (2004)	2/5.10-6.8/M-F/ASD (2)/IQ_F:54	Special education class/1:1	ATD	Mands, motor imitation and word vocalizations	MS, PECS	Progressive time delay procedure, PECS protocol, The most to least prompting procedure, model	–	–	Y/N/Y	Y/Y
van der Meer, Didden et al. (2012)	3/6-12-13/2 M, 2 F/ASD (1), PDD-NOS (1), CDD(1)	Therapy room/1:1	MPD with ATD	Request	MS, PE, SGD	Discrete trial teaching, 10 s. time delay procedure, graduated guidance	SGD(2), PE(1)	Researcher	N/Y/N	Y/Y
van der Meer, Kagohara et al. (2012)	2/7-10/2 M/Autism (2)	Special education class/1:1	MPD with ATD	Request	MS, SGD	Discrete trial teaching, 10 s. time delay procedure, graduated guidance	MS(1), SGD(1)	Graduate assistant	N/Y/N	Y/Y
van der Meer, Kagohara et al. (2013)	2/10-11/M-F/Autism (2)	Special education class (1), dining room at home(1)/1:1	ATD	Request, greeting, saying please and thanking, answering the question with yes-no	MS, PE, SGD	The least to most prompting procedure,	SGD(1), PE(1)	Mother, researcher at the beginning, then experienced practitioner	N/Y/N	Y/Y
van der Meer, Sutherland et al. (2012)	4/4-4-10-11/3M-1F/Autism (4)	The dining room at home, special education class/1:1	ATD	Request	MS, PE, SGD	Discrete trial teaching, 10 s. time delay procedure, graduated guidance	SGD(2), PE(2)	Mother, teacher assistant	N/Y/N	Y/Y

*: studies with the same participants; ATD: Alternating Treatment Design; CDD: Childhood Disintegrative Disorder; G: Generalization; IOA: Interobserver Agreement; IQ: Intelligence Quotes; M: Maintenance; MBD: Multiple Baseline Design; MPD: Multiple Probe Design; MS: Manual Sign; MT-RD: Multiple Treatment-Reversal Design; PECS: Picture Exchange Communication Systems; PDD-NOS: Pervasive Developmental Disorder Not Otherwise Specified; SGD: Speech Generating Device; SV: Social validity; TF: Treatment Fidelity; VOCA: Voice Output Communication Aid

room). Seven of the studies (e.g. van der Meer, Didden et al., 2012) were conducted in a therapy/practice/teaching room or in a quiet room, two of the studies (Boesch et al., 2013b; Couper et al., 2014) were carried out in a clinic, and one study (Scholosser et al., 1998) was conducted in the school library. In 16 studies, information about instructional arrangement was provided. The vast majority of the studies (n = 11; e.g. Son et al., 2006) were conducted as one-to-one instructional arrangement. In five studies (e.g. Schlosser & Blischak, 2004), more than one practitioner (two or three practitioners) intervened with the student in the instructional environment.

While the alternating treatment design was used in the majority of the studies (n = 12; e.g. van der Meer et al., 2012), in five studies (e.g. van der Meer, Kagohara et al., 2012), the alternating treatment design was used together with multiple probe/baseline design. Furthermore, the adaptive alternating treatment design was used in two studies (Schlosser et al., 1998; Schlosser & Blischak, 2004), the alternating treatment design was used together with the nonconcurrent multiple probe design in one study (McLay et al., 2017), and the multiple treatment with reversal design (ABCACB) was used in one study (Chen et al., 2016).

The request skill was studied in most of the studies (n = 18; e.g. Boesch et al., 2013a). While in three of the studies (e.g. Boesch et al., 2013b; van der Meer et al., 2013), interventions for social interaction (e.g. eye contact, physical orientation, etc.) and communication (e.g. greeting, answering the question, thanking, etc.) skills were conducted, in the other three studies (e.g. Beck et al., 2008), interventions aimed at speech production/vocalization were conducted, in two studies (Schlosser et al., 1998; Schlosser & Blischak, 2004), interventions for the spelling skills in the context of academic skills were conducted, and in one study (Tincani, 2004), interventions aimed at instructing imitation skills (motor imitation) were conducted. In five of the studies (e.g. Boesch et al., 2013b), interventions for more than one dependent variable were conducted.

Three applications were compared together in some of the studies when the studies were examined within the scope of the compared AAC applications (n = 10; e.g. van der Meer, Sutherland et al., 2012). While applications for manual sign, PE/PECS (Picture Exchange/Picture Exchange Communication Systems), and speech generating devices (SGD) or Voice Output Communication Aid (VOCA) were compared in seven of these studies (e.g. Achmadi et al., 2014), different modelings (e.g. print-speech-print+speech) used in SGD were compared in three of the studies (e.g. Schlosser & Blischak, 2004). In another part of the studies (n = 10; e.g. Lorah et al., 2013), two applications were compared together. In seven of these studies (e.g. Son et al., 2006), applications performed with speech generating devices (VOCA, SGD or SGD with the iPad) were compared with the PECS. In Tincani's (2004) study, manual sign and the PECS; in van der Meer, Kagohara, et al.'s (2012) study manual sign and SGD; in Chen, et al.'s (2016) study, SGD applications with two different interfaces (pie abbreviation/expansion - hierarchical relating) were compared. In one of the studies (Gevarter et al., 2017), SGD applications with four different modelings (photo image- symbol grid - hybrid – pop-up symbol grid) were compared together. In most of the studies (n = 20), various instructional contents were added to the teaching process while AAC applications were used. The time delay procedure (n = 12; e.g. McLay et al., 2015), graduated guidance (n = 6; e.g. McLay et al., 2017), the least or the most prompting procedure (n = 5; e.g. Gevarter et al., 2017), discrete-trial teaching (n = 5; e.g. Couper et al., 2014), PECS protocol (n = 5; e.g. Aguis & Vance, 2016), and practices such as prompting or being a model, error correction and reinforcement were used in the studies. Moreover, in two studies (Schlosser & Blischak, 2004; Schlosser et al., 1998), the Copy-Cover-Compare application was conducted in relation to the dependent variable (spelling skill).

Information related to the evaluation about preferences of the participants was obtained from 14 studies (e.g. Agius et al., 2016). In these studies, 35 participants (e.g. Couper et al., 2014) preferred SGD applications, seven participants (e.g. Son et al., 2006) preferred the PECS, and one participant (van der Meer, Kagohara et al., 2012)

preferred manual sign. One of the participants (Couper et al., 2014) did not prefer any application. In the preference evaluation of the comparative studies on SGD in different modelings (n = 2; Schlosser & Blischak, 2004; Schlosser et al., 1998), four participants preferred visual SGD, and one participant preferred audio SGD.

In most of the studies (n = 16; e.g. Son et al., 2006), information about practitioners was provided. The studies (n = 10; e.g. Gevarter et al., 2017) which involved more than one practitioner in an intervention constituted the majority of them, on the other hand some of them (n = 6; e.g. Aguis & Vance, 2016) conducted with a single practitioner. Most of the studies (n = 10; e.g. Achmadi et al., 2014) were conducted by researchers/research assistants. Furthermore, special education teachers or assistant teachers (n = 6; e.g. Lorah, 2016), graduate students (n = 3; e.g. Lorah et al., 2013), mothers (n = 2; van der Meer et al., 2013; van der Meer, Sutherland et al., 2012), therapist/child psychologist (Couper et al., 2014) and speech-language pathologist (McLay et al., 2015) were involved in the studies as practitioners.

In most of the studies (n = 17; e.g. Chen et al., 2016), generalization data were not collected. The studies (n = 4) in which generalization data were obtained from settings, practitioners, participants or materials. Maintenance data were obtained nearly in half of the studies (n = 12; e.g. Boesch et al., 2013a). While three studies were obtained maintenance data after criteria (Lorah et al., 2013; Scholosser & Blischak, 2004; Schlosser et al., 1998), in other studies (n = 9) the maintenance data were collected at times after two weeks to seven months. Social validity data were obtained from teachers/teacher assistants/therapists (n = 5; e.g. Lorah, 2016) and parents/caregivers (n = 3; e.g. Aguis et al., 2016). Interobserver agreement data were obtained in all studies (n = 20) except for one study (Chen et al., 2016). The reliability coefficients vary between 80% and 100%. Treatment fidelity was obtained in all studies except for two studies (Chen et al., 2016; Son et al., 2006), the treatment fidelity coefficients vary between 89% and 100%.

Results of Quality Indicators

The evaluation of the included studies in terms of quality indicators is presented in Table 2. Since Horner et al. published quality indicators at 2005, the studies were examined with respect to before and after 2005. There were three studies (Scholosser et al., 1998; Schlosser & Bischak, 2004; Tincani, 2004) published before 2005. Among these studies, no study met all of the quality indicators. However, all of these studies (n = 3) were coded as *"Y (Yes)"* in terms of the following indicators: adequate identification of participant characteristics (item 1), adequate identification of the participant selection process (item 2), all indicators in the category of the dependent and independent variable and baseline (items between 4–13), the dependent variable being socially important (item 18), and implementation of the independent variable over extended time in typical (natural) contexts/persons (item 21). While two of the studies (Schlosser & Bischak, 2004; Tincani, 2004) were coded as *"Y (Yes)"* in terms of items 14, 15, 16, and 17 in the validity category, the other study (Schlosser et al., 1998) was coded as *"N (No)"* in terms of these indicators. While items 19 and 20 among the quality indicators were coded as *"Y (Yes)"* in two studies (Schlosser et al., 1998, Tincani, 2004), in the other study (Schlosser & Bischak, 2004), they were coded as *"N (No)."* Furthermore, the indicator *"the setting was defined adequately"* (item 3) was coded as *"N (No)"* in all three studies.

The majority of the studies (n = 18) were published after 2005. There were also none of them that met all quality indicators. However, all of them were coded as *"Y (Yes)"* in terms of the following indicators: adequate identification of participant characteristics (item 1), performing repeated measurements regarding the dependent variable (item 7), being described with replicable precision regarding the independent variable, and being systematically manipulated (items 9 and 10), and socially important dependent variable (item 18). While all of the studies (n = 17) except for one study (Chen et al., 2016) were coded as *"Y (Yes)"* in terms of reporting the interobserver agreement data (item 8), all other studies (n = 16), except for two studies (Chen et al., 2016; Son et al., 2006), were coded as *"Y (Yes)"*

TABLE 2

Evaluation of the Studies in Terms of the "Quality Indicators of Single-Subject Experimental Research Studies"

Quality Indicators	*Achmadi et al. (2014)*	*Aguis and Vance (2016)*	*Beck et al. (2008)*	*Boesch et al. (2013a)*	*Boesch et al. (2013b)*	*Chen et al. (2016)*	*Couper et al. (2014)*	*Gevarter et al. (2017)*	*Gevarter et al. (2014)*	*Lorah (2016)*
Participants and Settings										
1. The participants were described adequately	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. The selection process was described adequately	N	Y	Y	Y	N	N	Y	N	Y	Y
3. The setting was described adequately	N	Y	N	Y	N	N	N	N	N	Y
Dependent Variable										
4. Described with operational precision	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
5. Measurable	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
6. The measurement was defined with replicable precision	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
7. Repetitive measurements were made	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8. Interobserver agreement data were reported	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Independent variable										
9. Described with replicable precision*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10. Systematically manipulated	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
11. Treatment fidelity was defined*	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Baseline										
12. Phase provided evidence for the design (pattern) before the application	N	Y	N	Y	Y	Y	Y	N	N	N
13. Described with replicable precision	Y	Y	Y	Y	Y	Y	Y	N	N	N
Validity										
14. The experimental effect has three proofs/representations*	Y	Y	Y	Y	Y	Y	Y	N	N	N
15. The design controlled the threats for internal validity*	Y	Y	N	Y	Y	Y	Y	N	N	N
16. The results presented a pattern that experimental control was demonstrated*	Y	Y	Y	Y	N	Y	Y	N	N	N
17. The experimental effects were replicated, external validity was provided*	Y	Y	Y	Y	Y	Y	Y	N	N	N
18. The dependent variable is socially important	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
19. The magnitude of change in the dependent variable resulting from the application is socially important	N	Y	N	Y	N	N	N	N	N	Y
20. The independent variable is cost effective and/or practical	N	Y	N	Y	N	N	N	N	N	Y
21. The independent variable was applied over extended time in the presence of typical (natural) contexts/persons	Y	N	Y	N	N	N	Y	Y	Y	Y
Indicators that are met/total indicators	**16/21**	**20/21**	**16/21**	**20/21**	**15/21**	**11/21**	**18/21**	**11/21**	**12/21**	**15/21**

TABLE 2—(Continued)

Quality Indicators	*Lorah et al. (2013)*	*McLay et al. (2015)*	*McLay et al. (2017)*	*Schlosser & Bischak (2004)*	*Schlosser et al. (1998)*	*Son et al. (2006)*	*Tincani (2004)*	*van der M., Didden et al. (2012)*	*van der M., Kagohara et al. (2012)*	*van der M., Kagohara, et al. (2013)*	*van der M., Sutherland et al. (2012)*
Participants and Settings											
1. The participants were described adequately	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. The selection process was described adequately	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3. The setting was described adequately	Y	N	N	N	N	N	N	N	Y	Y	N
Dependent Variable											
4. Described with operational precision	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
5. Measurable	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
6. The measurement was defined with replicable precision	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
7. Repetitive measurements were made	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8. Interobserver agreement data were reported	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Independent variable											
9. Described with replicable precision*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10. Systematically manipulated	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
11. Treatment fidelity was defined*	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Baseline											
12. Phase provided evidence for the design (pattern) before the application	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
13. Described with replicable precision	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Validity											
14. The experimental effect has three proofs/representations*	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y
15. The design controlled the threats for internal validity*	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y
16. The results presented a pattern that experimental control was demonstrated*	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y
17. The experimental effects were replicated, external validity was provided*	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y

TABLE 2—(Continued)

Quality Indicators	*Lorah et al. (2013)*	*McLay et al. (2015)*	*McLay et al. (2017)*	*Schlosser & Bischak (2004)*	*Schlosser et al. (1998)*	*Son et al. (2006)*	*Tincani (2004)*	*van der M., Didden et al. (2012)*	*van der M., Kagohara et al. (2012)*	*van der M., Kagohara,Sutherland et al. (2013)*	*van der M., et al. (2012)*
18. The dependent variable is socially important	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
19. The magnitude of change in the dependent variable resulting from the application is socially important	N	N	N	N	Y	N	Y	N	N	N	N
20. The independent variable is cost effective and/or practical	N	N	N	N	Y	N	Y	N	N	N	N
21. The independent variable was applied over extended time in the presence of typical (natural) contexts/persons	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y
Indicators that are met/total indicators	**18/21**	**18/21**	**10/21**	**18/21**	**16/21**	**17/21**	**20/21**	**17/21**	**19/21**	**14/21**	**18/21**

in terms of defining the treatment fidelity (item 11), and all of the other studies (n = 15), except for three studies (Chen et al., 2014; Gevarter et al., 2017; Gevarter et al., 2014), were coded as *"Y (Yes)"* in terms of being defined with replicable precision regarding the baseline (item 13). Most of the studies (n = 16, e.g. Son et al., 2006) were coded as *"Y (Yes)"* in terms of items 4, 5 and 6 in the dependent variable category. Besides, most of the studies (n = 13; e.g. Aguis & Vance, 2016) were coded as *"Y (Yes)"* in terms of the adequate identification of the participant selection process (item 2). Furthermore, 12 studies (e.g. Couper et al., 2014) in terms of the indicator of phase provided evidence for the design (pattern) before the application (item 12), other 12 study groups (e.g. Gevarter et al., 2014) in terms of the indicator of implementing the independent variable over time in typical (natural) contexts/persons (item 21), and other study groups (n=12; e.g. Achmadi et al., 2014) in terms of the indicators of having three representations of the experimental effect, controlled threats related to the internal validity of the design, and repeating the effect and providing external validity (items 14, 15 and 17, respectively) were coded as *"Y (Yes)"*. 11 studies (e.g. van der Meer, Sutherland et al., 2012) were coded as *"Y (Yes)"* in terms of the indicator (item 16) of providing a pattern that experimental control is demonstrated. A small number of the studies (n = 6; e.g. McLay et al., 2017) were coded as *"Y (Yes)"* with regard to the indicator of identifying the setting adequately (item 3). However, only three studies (Aguis & Vance, 2016; Boesch et al., 2013a; Lorah, 2016) were coded as *"Y (Yes)"* in terms of the following two indicators in the validity category: "depending on the application, the size of the change in the dependent variable is socially important (item 19)" and "the independent variable is cost effective and/or practical (item 20)."

The studies were evaluated in terms of the *"acceptability"* criteria as studies published before and after 2005. According to this, items 9, 11, 14, 15, 16 and 17 among quality indicators have been considered in terms of the *"acceptability"* criteria. In this context, while two of the studies published before 2005 (Schlosser & Bischak, 2004; Tincani, 2004) were evaluated as *"acceptable"*, nine of the studies published after 2005 (e.g. Aguis & Vance, 2016; Lorah et al., 2013) were evaluated as *"acceptable"* studies. As a result, among the included studies (n = 21), 11 studies were evaluated as *"acceptable"* studies. These studies which were evaluated as *"acceptable"* were filed for inclusion in the WWC design standards and meta-analysis processes.

Results of Effect Size and WWC (2017) Design Standards

Table 3 presents the evaluation results of the studies in terms of the design standards and meta-analysis findings. In terms of the design standards, only one of the studies (Boesch et al., 2013a) meets the design standards. In four of the studies (Aguis & Vance, 2016; Lorah et al., 2013; Tincani, 2004; VanderMeer, Kagohara et al., 2012), since the criterion of having five and more data points for per condition is not met in one phase but it is met in other phases, they meet design standards with reservation. Because more than two conditions are compared in four of the studies (Couper et al., 2014; McLay et al., 2015; Schlosser & Blischak, 2004; van der Meer, Didden et al., 2012; van der Meer, Sutherland, 2012), the carryover effect of the applications cannot be controlled, and so they do not meet the WWC design standards. In one of the studies (Achmadi et al., 2004), data were obtained two or three times in all of the baseline phases, so it was evaluated as not meeting design standards.

In terms of IRD effect size calculations, eight of the studies (e.g. Achmadi et al., 2014) which were implemented with SGD are *high effect* applications in teaching skills such as request/mands and communication skills by being evaluated over .70. In one of the studies (Couper et al., 2014), the effect size of SGD application in the teaching of the request skill was calculated to be 95% CI (.47, 86) .68, and evaluated as a *medium effect* application. The IRD effect size in all of the studies (n = 9; e.g. Tincani et al., 2004) in which picture exchange-based teaching (PE/PECS) were compared with other applications (SGD, manual

TABLE 3

Evaluation of the Studies in Terms of the WWC Design Standards and 95% CI IRD Effect Size Findings

				95% CI	
Study	*Research Design*	*WWC (2017)*	*IRD (Mean)*	*Low*	*High*
Achmadi et al. (2014)	ATD	Not meet	IRD_{SGD}: 1	1.00	1.00
			IRD_{PE}: .91	.83	.91
			IRD_{MS}: .62	.47	.76
*Aguis & Vance (2016)	MBD with ATD	Meets with reservation	IRD_{SGD}: .95	.90	.95
			IRD_{PECS}: .96	.93	.96
*Boesch et al. (2013a)	MBD with ATD	Meets	IRD_{SGD}: .96	.93	.98
			IRD_{PECS}: .91	.82	.91
Couper et al. (2014)	ATD	Not meet (For three cases)	IRD_{SGD}: .68	.47	.86
			IRD_{PE}: .81	.70	.91
			IRD_{MS}: .37	.14	.63
*Lorah et al. (2013)	ATD	Meets with reservation	IRD_{SGD}: 1	1.00	1.00
			IRD_{PE}: .98	.94	.98
McLay et al. (2015)	ATD	Not meet (For three applications)	IRD_{SGD}: .86	.73	.97
			IRD_{PE}: .72	.63	.80
			IRD_{MS}: .40	.11	.63
Schlosser & Blischak (2004)	ATD	Not meet (For three applications)	IRD_{PRINT}: .78	.67	.89
			IRD_{SPEECH}: .79	.68	.89
			$IRD_{PRINT\text{-}SPE.}$: .65	.34	.89
*Tincani (2004)	ATD	Meets with reservation	IRD_{PECS}: .93	.79	.93
			IRD_{MS}: .91	.82	.97
van der Meer, Didden et al. (2012)	ATD	Not meet (For three applications)	IRD_{SGD}: .89	.81	.96
			IRD_{PE}: .97	.94	.97
			IRD_{MS}: .47	.18	.76
*van der Meer, Kagohara et al. (2012)	MPD with ATD	Meets with reservation	IRD_{SGD}: .87	.87	.87
			IRD_{MS}: .57	.57	.57
van der Meer, Sutherland et al. (2012)	ATD	Not meet (For three applications)	IRD_{SGD}: .85	.61	.98
			IRD_{PE}: .95	.87	.95
			IRD_{MS}: .70	.25	.95

***:** *Acceptable studies in terms of the WWC standards*

sign) in the teaching of skills such as requesting, vocalization, motor imitation was calculated to be higher than .70, so evaluated as a *high effect*. While three of the applications (Couper et al., 2014; McLay et al., 2015; van der Meer, Didden et al., 2012) performed with manual sign were considered to have a *small effect* compared to other applications (PECS or SGD) by being evaluated under .50, three applications (Achmadi et al., 2014; van der Meer, Kagohara et al., 2012; van der Meer, Sutherland et al., 2012) were estimated to have a *medium effect size* in teaching request skills compared to other applications by calculating the IRD effect size between .50 and .70. In one of the studies (Tincani, 2004), the application of manual sign was evaluated as an application with a *high effect* by calculating the IRD effect size as 95% (.82, .97) .91.

In another study (Schlosser & Blischak, 2004), in contrast to communication skills, the device generating three different speeches (print, speech, print + speech) was compared in the teaching of the skill of spelling among academic skills. Among the compared applications, SGD print and speech applications as separately were calculated to be higher than .70, and they were evaluated as *high effect* in the teaching of the spelling skill. However, the effect size of SGD both print and speech application was calculated to be 95% CI (.34, .89) .65, and it was evaluated as *medium effect*.

Discussion

In summary, while PECS and SGD applications among AAC applications are resulted to be equally effective in teaching social communication skills, the manual sign application is not effective at the desired level. Furthermore, there are too many variables (e.g. participant characteristics, instructional content) which may change the effectiveness of AAC applications. In addition to these findings, a remarkable finding that may affect credibility in experimental processes is that most of the studies were evaluated as not meeting design standards in terms of the WWC design standards since they compared three different applications in the same phase. Although it is not recommended to calculate the effect size of studies not meeting the WWC design standards, the rubric of Horner et al. was preferred in this article, therefore the effect sizes of the studies assessed as "acceptable" were calculated. However, it was surprisingly resulted that almost half of the studies assessed as "acceptable" could not meet the WWC design standards.

A large number of participants in the studies comparing AAC applications were either in the pre-school period or in the primary school period. The reason for the fact that the participants in the studies were mostly selected from these age ranges should be a result of considering the importance of early intervention in language acquisition. Although there were studies comparing various applications in secondary school students, there was not any study comparing AAC applications in the young/adulthood period. However, in the literature, there are a limited number of studies (e.g. Kagohara et al., 2010, Lund & Troha, 2008) which examine the effectiveness of AAC applications aimed to teach such as request/mands and answering questions to individuals who are over or at the young age (e.g. 17 years) and have ASD. Nevertheless, most of adult ASD have continued limited communication skills.

When the studies were reviewed in terms of the setting in which they were conducted, we realised that the researchers mostly preferred settings in natural contexts (e.g. student's class, home environment). This situation is important according to Horner et al. in ensuring that the quality indicator, which is "*the implementation of the independent variable in typical contexts over extended time"* is met. In all of the studies except for one study (Chen et al., 2016), the alternating treatment design (with adapted in some studies or together with the multiple probe/baseline designs) was used. It can be stated that the reason for preferring mostly the alternating treatment design is that this design can be completed as soon as possible among the comparative designs, and it is the most suitable design for controlling the factors affecting internal validity. A remarkable finding is that more than two applications are compared alternately in almost half of the studies. This situation raises the issue of carryover effects related to the dependent variable depending on the applications. The situation accepted in terms of the WWC (2017) design standards is that this effect can be minimized by comparing two applications within the alternating treatment design.

The skill of request/mands is the most targeted skill among the behaviors intended to be taught to participants with AAC applications in the studies. In other review studies, which examine the effectiveness of AAC applications conducted for individuals with ASD in the literature (e.g. Ganz et al., 2011; Lorah et al., 2015; van der Meer & Rispoli, 2010), it is seen that the request skill was studied intensively individuals with ASD, and besides this skill, other skills such as social skills, academic skills, behavior problems were also studied in some studies. According to van der Meer and Rispoli (2010), it is not surprising that there is a high tendency to teach this skill in most AAC application studies conducted for these individuals, since individuals with ASD do not acquire request skills with a natural process. Furthermore, although the request is a pivotal communication skill, it may be appropriate to acquire it before teaching more complex communication skills.

In the studies, mostly manual sign, PECS or SGD applications were compared with each other in the teaching of skills such as requesting, speech production, and vocalization. Although these applications seem like a single application, various contents such as time delay procedure, graduated guidance, the least to most/the most to least prompting procedure, discrete trial teaching have been incor-

parated into applications. Therefore, it can be said that various errorless teaching methods can be used to increase the effectiveness of AAC applications and to instruct how to use them.

Mostly participants who particapeted the studies have preferred SGD applications. This is an indication that individuals with ASD are more interested in technologically based applications. At the same time, it supports the fact that there is an increasing tendency, in the literature, to technologically based applications in teaching various skills to individuals with ASD (Grynszpan et al., 2014).

Generalization and social validity data are obtained from a limited number of studies. Although this situation is expressed as a limitation in the studies, the data can not be obtained because it can make the research process more difficult since there are more than one applications in the comparative studies. In future studies comparing AAC applications, obtaining generalization and social validity data from only effective or preferred application is an issue that should be focused on by researchers because this approach can ensure both the easy planning of the process and eliminating the negative tendencies exhibited by individuals with ASD in generalization and maintenance skills. Maintenance data were obtained from half of the studies. Long-term (e.g. 5-month, 7-month) maintenance data is important to demonstrate that the long-term effects of AAC applications can be sustained.

Interobserver agreement data and treatment fidelity data were obtained in most of the studies. The mentioned data were also obtained for most of the studies in which AAC applications with different inclusion criteria were systematically reviewed. The fact that especially the treatment fidelity data were obtained in the studies is considered important regarding to meet the criterion of treatment fidelity in evaluating a study as "*acceptable"* in terms of the *"Quality Indicators of Single-Subject Experimental Research Studies."*

There was no significant difference in the analysis results regarding the quality indicators of the included studies in terms of being published before or after 2005. However, the number of studies published before 2005 was limited. Therefore, there are no findings suitable for comparison. However, it is primarily determined that the quality indicator related to defining the setting adequately was not met in most of the studies. Furthermore, due to obtaining social validity data in a limited number of studies, the indicators of *"the change in the dependent variable due to the application being socially significant"* and *"the independent variable being cost-effective or practical"* were evaluated as *"No"* in most of the studies. The studies evaluated as *"acceptable"* are about half of the included studies. In the other studies which were evaluated as not *"acceptable,"* because the items (14, 15, 16, and 17) evaluated as "acceptable" in the validity category are not met. In addition, the fact that the baseline data were not obtained caused the studies to be coded as *"No"* in terms of these indicators. The reason for this is that although obtaining baseline data is not a required condition in alternating treatment designs, it is recommended in order to see the experimental effect (Horner et al., 2005). Furthermore, not obtaining the baseline data is especially recommended for behaviors that need to be reduced.

In terms of research designs, the number of studies meeting the WWC (2017) design standards was quite few. Most of the studies did not meet the WWC design standards since more than two applications were compared within the phase. Besides, five or more data points in each phase is a rigorous criterion in the studies, and in most of the studies, five and more data points were not obtained in per phase, especially at the baseline. Therefore, the studies meeting the design standards met them with reservation, except for one study (Boesch et al., 2013a).

When AAC applications are compared regarding the IRD effect size, the most ineffective applications are the manual sign applications. This is due to the inadequacies in the motor skills of children with ASD, as well as the difficult and complex structure of the manual sign application compared to other AAC applications (Mirenda, 2003; van der Meer, Didden et al., 2012). However, the manual sign application may be a more effective application than aided AAC applications (e.g. PECS) in acquiring communication skills for individuals with ASD who have high-performance levels regarding motor skills (Tincani, 2004).

SGD and PECS applications are effective in

teaching basic communication skills (requesting, short answering, etc.) to individuals with ASD. In the teaching of communication skills to these individuals, practitioners may be recommended to prefer these applications, but it may be also recommended to consider preferences of these individuals. In the literature, it is indicated that individuals with ASD prefer SGD applications more. As recommended by Muharib and Alzrayer (2018), high-tech SGD applications such as tablets and iPads are suitable for use in classroom environments. Most of the applications conducted with the manual sign are ineffective or have a moderate effect and are not preferred. Besides, it can be stated that it is an effective and preferred application on individuals with ASD who have high motor performance levels. Therefore, it may be recommended for practitioners to obtain comprehensive information of student characteristics for determining the application. Furthermore, the studies conducted with AAC applications show that various evidence-based applications such as error correction, reinforcement, discrete trial teaching, and time delay procedure are used within the process. Therefore, it may be recommended to practitioners who will use these applications to benefit from evidence-based practices in order to make the process more efficient. In the literature, the request skill in the context of communication skills to individuals with ASD has preferred more intensively. If the fact that these individuals exhibit behavioral problems in case they are unable to express themselves is considered (Chiang, 2008; Park et al., 2012), it may be suggested that the teaching of the skill of request/mands should be preferred as a priority by practitioners.

Implications for the Future Research

In the future comparative studies, it may be suggested to design a study in such a way that maximum two AAC applications will be compared within the per condition. If more than two AAC applications are to be compared in a study, research should be designed to compare each application with each other separately (e.g. comparison of manual sign and PECS, comparison of PECS and SGD, comparison of SGD and manual sign). In a large part of the studies designed according to the alternating treatment design, the process was conducted without obtaining the baseline data. According to Horner et al. (2005), there should be a significant difference between the baseline and intervention phases in order to have evidence of the experimental effect. Therefore, there is a suspicion that this evidence has been provided in the studies, in which baseline data have not been obtained. As a result of this, the studies are evaluated negatively in terms of the *"acceptability"* criteria. Besides, in terms of the WWC design standards, it is recommended to obtain at least five data points in each phase (baseline, intervention). Therefore, it is recommended to arrange baseline phases and to obtain at least five data points for per phase in comparative studies which will be conducted in the future.

In most of the studies, characteristics of the setting are not described in such a way as to enable the repetition. In future studies, it may be suggested to increase the repeatability of studies by describing the setting characteristics comprehensively. Social validity data and generalization data were not obtained in most of the studies. However, in comparative studies, which one of the applications is more effective and efficient can be determined by taking opinions from the parents, teachers or students. Furthermore, ability of students to use the application in different settings or in the presence of different individuals is important data for the generalization of the study results. On the other hand, the impact precision of SGD applications is low. There is a need for new studies to determine which one of these applications is more effective in teaching different skills to different participants in different conditions.

In conclusion, in this descripitve and meta-analysis study, we realised that teaching the request is mostly aimed in AAC application studies conducted on individuals with ASD. However, different skills are also taught. PECS and SGD applications come to the forefront as highly effective applications. Nevertheless, a large number of individuals with ASD prefer SGD applications. In descriptive and meta-analysis studies to be conducted in the future, it may be suggested to evaluate studies by analyzing different variables (e.g. efficiency, effect sizes of other data) with different analysis techniques (e.g. the percentage of non-

overlapping data, Tau-*U*). The WWC (2017) recommends calculating the effect size of studies that meet the design standards with or without reservation. In this study, the effect size of the studies evaluated as *"acceptable"* recommended by Horner et al. (2005) was calculated since the number of the studies meeting the design standards with or without reservation was limited. In meta-analysis studies to be conducted in the future, only the effect sizes of studies which meet the design standards (with/without reservation) may be calculated.

References

*Included Studies

**Excluded Studies

*Achmadi, D., Sigafoos, J., van der Meer, L., Sutherland, D., Lancioni, G. E., O'Reilly, M. F., & Marschik, P. B. (2014). Acquisition, preference, and follow-up data on the use of three AAC options by four boys with developmental disability/delay. *Journal of Developmental and Physical Disabilities, 26,* 565–583.

*Agius, M. M., & Vance, M. (2016). A comparison of PECS and iPad to teach requesting to preschoolers with autistic spectrum disorders. *Augmentative and Alternative Communication, 32,* 58–68.

American Psychiatric Association. (2013). *Ruhsal bozuklukların tanısal ve sayımsal el kitabı: Beşinci Baskı (DSM-5*) [*Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition (DSM-5*)]. (Çev: E. KörogĬu). Ankara: Hekimler Yayın Birligĭ.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd Ed.). Author.

Aydın, O., Tekin-İftar, E., & Rakap, S. (2019). Explanation of "quality indicators of single-subject experimental research studies" rubric using studies focused on teaching mathematics skills. *Ankara University Faculty of Educational Sciences Journal of Special Education, 20,* 597–628

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z.,. . . Durkin, M. S. (2018). Prevalence of autism spectrum disorder among children aged 8 years. Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveillance Summaries, 67,* 1–23.

*Beck, A. R., Stoner, J. B., Bock, S. J., & Parton, T. (2008). Comparison of PECS and the use of a VOCA: A replication. *Education and Training in Developmental Disabilities, 43,* 198–216.

*Boesch, M. C., Wendt, O., Subramanian, A., & Hsu, N. (2013a). Comparative efficacy of the Picture Exchange Communication System (PECS) versus a speech-generating device: Effects on requesting skills. *Research in Autism Spectrum Disorders, 7,* 480–493.

*Boesch, M. C., Wendt, O., Subramanian, A., & Hsu, N. (2013b). Comparative efficacy of the Picture Exchange Communication System (PECS) versus a speech-generating device: effects on social-communicative skills and speech development. *Augmentative and Alternative Communication, 29,* 197–209.

Bondy, A., & Frost, L. (1994). The Picture Exchange Communication System. *Focus on Autistic Behavior, 9,* 1–19.

Boyd, T. K., Hart Barnett, J. E., & More, C. M. (2015). Evaluating iPad technology for enhancing communication skills of children with Autism Spectrum Disorders. *Intervention in School and Clinic, 51,* 19–27.

Boyle, M. A., Samaha, A. L., Rodewald, A. M., & Hoffmann, A. N. (2013). Evaluation of the reliability and validity of GraphClick as a data extraction program. *Computers in Human Behavior, 29,* 1023–1027.

Carroll, R. A., & Kodak, T. (2014). An evaluation of interrupted and uninterrupted measurement of vocal stereotypy on perceived treatment outcomes. *Journal of Applied Behavior Analysis, 47,* 264–276.

Centers for Disease Control and Prevention. (2014). Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010. Morbidity and mortality weekly report. *MMWR Surveillance Summaries, 63,* 1–21.

*Chen, C. H., Wang, C. P., Lee, I. J., & Su, C. C. C. (2016). Speech-generating devices: effectiveness of interface design a comparative study of autism spectrum disorders. *SpringerPlus, 5,* 1682.

Chiang, H. M. (2008). Expressive communication of children with autism: The use of challenging behaviour. *Journal of Intellectual Disability Research, 52,* 966–972.

*Couper, L., van der Meer, L., Schäfer, M. C., McKenzie, E., McLay, L., O'Reilly, M. F., . . . Sutherland, D. (2014). Comparing acquisition of and preference for manual signs, picture exchange, and speech-generating devices in nine children with autism spectrum disorder. *Developmental Neurorehabilitation, 17,* 99–109.

Fletcher-Watson, S., Petrou, A., Scott-Barrett, J., Dicks, P., Graham, C., O'Hare, A., . . . McConachie, H. (2016). A trial of an iPad™ intervention targeting social communication skills in children with autism. *Autism, 20,* 771–782.

Ganz, J. B., Davis, J. L., Lund, E. M., Goodwyn, F. D., & Simpson, R. L. (2012). Meta-analysis of PECS with individuals with ASD: Investigation of targeted versus non-targeted outcomes, participant

characteristics, and implementation phase. *Research in Developmental Disabilities, 33,* 406–418.

Ganz, J. B., Earles-Vollrath, T. L., Heath, A. K., Parker, R. I., Rispoli, M. J., & Duran, J. B. (2012). A meta-analysis of single case research studies on aided augmentative and alternative communication systems with individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42,* 60–74.

Ganz, J. B., Earles-Vollrath, T. L., Mason, R. A., Rispoli, M. J., Heath, A. K., & Parker, R. I. (2011). An aggregate study of single-case research involving aided AAC: Participant characteristics of individuals with autism spectrum disorders. *Research in Autism Spectrum Disorders, 5,* 1500–1509.

**Ganz, J. B., Hong, E. R., Gilliland, W., Morin, K., & Svenkerud, N. (2015). Comparison between visual scene displays and exchange-based communication in augmentative and alternative communication for children with ASD. *Research in Autism Spectrum Disorders, 11,* 27–41.

*Gevarter, C., O'Reilly, M. F., Kuhn, M., Watkins, L., Ferguson, R., Sammarco, N., . . . Sigafoos, J. (2017). Assessing the acquisition of requesting a variety of preferred items using different speech generating device formats for children with autism spectrum disorder. *Assistive Technology, 29,* 153–160.

*Gevarter, C., O'Reilly, M. F., Rojeski, L., Sammarco, N., Sigafoos, J., Lancioni, G. E., & Lang, R. (2014). Comparing acquisition of AAC-based mands in three young children with autism spectrum disorder using iPad® applications with different display and design elements. *Journal of Autism and Developmental Disorders, 44,* 2464–2474.

Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism, 18,* 346–361.

Hart, S. L., & Banda, D. R. (2010). Picture Exchange Communication System with individuals with developmental disabilities: A meta-analysis of single subject studies. *Remedial and Special Education, 31,* 476–488.

**Hill, D. A., & Flores, M. M. (2014). Comparing the Picture Exchange Communication System and the iPad (TM) for Communication of Students with Autism Spectrum Disorder and Developmental Delay. *TechTrends, 58,* 45–53.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71,* 165–179.

Kagohara, D. M., van der Meer, L., Achmadi, D., Green, V. A., O'Reilly, M. F., Mulloy, A., . . . Sigafoos, J. (2010). Behavioral intervention promotes successful use of an ipod-based communication device by an adolescent with autism. *Clinical Case Studies, 9,* 328–338.

Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., . . . Almirall, D. (2014). Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 53,* 635–646.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & ShAACh, W. R. (2010). Single-case designs technical documentation Retrieved from: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Logan, K., Iacono, T., & Trembath, D. (2017). A systematic review of research into aided AAC to increase social-communication functions in children with autism spectrum disorder. *Augmentative and Alternative Communication, 33,* 51–64.

*Lorah, E. R. (2016). Comparing Teacher and Student Use and Preference of Two Methods of Augmentative and Alternative Communication: Picture Exchange and a Speech-Generating Device. *Journal of Developmental and Physical Disabilities, 28,* 751–767.

Lorah, E. R., Parnell, A., Whitby, P. S., & Hantula, D. (2015). A systematic review of tablet computers and portable media players as speech generating devices for individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45,* 3792–3804.

*Lorah, E. R., Tincani, M., Dodge, J., Gilroy, S., Hickey, A., & Hantula, D. (2013). Evaluating picture exchange and the iPad™ as a speech generating device to teach communication to young children with autism. *Journal of Developmental and Physical Disabilities, 25,* 637–649.

Lund, S. K., & Troha, J. M. (2008). Teaching young people who are blind and have autism to make requests using a variation on the Picture Exchange Communication System with tactile symbols: A preliminary investigation. J*ournal of Autism and Developmental Disorders, 38,* 719–730.

*McLay, L., Schäfer, M. C., van der Meer, L., Couper, L., McKenzie, E., O'Reilly, M. F., . . . Sutherland, D. (2017). Acquisition, Preference and Follow-up Comparison Across Three AAC Modalities Taught to Two Children with Autism Spectrum Disorder. *International Journal of Disability, Development and Education, 64,* 117–130.

*McLay, L., van der Meer, L., Schäfer, M. C., Couper, L., McKenzie, E., O'Reilly, M. F., . . . Sutherland, D. (2015). Comparing acquisition, generalization, maintenance, and preference across three AAC options in four children with autism spectrum disorder. *Journal of Developmental and Physical Disabilities, 27,* 323–339.

Mirenda, P. (2003). Toward functional augmentative and alternative communication for students

with autism: Manual signs, graphic symbols, and voice output communication aids. *Language, Speech, and Hearing Services in Schools, 34,* 203–216.

Mueller, V. T. (2014). Alternative communication. In F. Volkmar, S. Rogers, R. Paul, & K. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 813–825). Wiley.

Muharib, R., & Alzrayer, N. M. (2018). The use of high-tech speech-generating devices as an evidence-based practice for children with autism spectrum disorders: A meta-analysis. *Review Journal of Autism and Developmental Disorders, 5,* 43–57.

O'Malley, P., Lewis, M. E. B., Donehower, C., & Stone, D. (2014). Effectiveness of using ipads to increase academic task completion by students with autism. *Universal Journal of Educational Research, 2,* 90–97.

Park, C. J., Yelland, G. W., Taffe, J. R., & Gray, K. M. (2012). Brief report: The relationship between language skills, adaptive behavior, and emotional and behavior problems in preschoolers with autism. *Journal of Autism and Developmental Disorders, 42,* 2761– 2766.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75,* 135–150.

Rakap, S., Rakap, S., Evran, D., & Cig, O. (2016). Comparative evaluation of the reliability and validity of three data extraction programs: UnGraph, GraphClick, and DigitizeIt. *Computers in Human Behavior, 55,* 159–166.

Rice, C. E., Rosanoff, M., Dawson, G., Durkin, M. S., Croen, L. A., Singer, A., & Yeargin-Allsopp, M. (2012). Evaluating changes in the prevalence of the autism spectrum disorders (ASDs). *Public Health Reviews, 34,* 1–22.

*Schlosser, R. W., Blischak, D. M., Belfiore, P. J., Bartley, C., & Barnett, N. (1998). Effects of synthetic speech output and orthographic feedback on spelling in a student with autism: A preliminary study. *Journal of Autism and Developmental Disorders, 28,* 309–319.

*Schlosser, R. W., & Blischak, D. M. (2004). Effects of speech and print feedback on spelling by children with autism. *Journal of Speech, Language, and Hearing Research, 47,* 848–862.

**Schlosser, R. W., Sigafoos, J., Luiselli, J. K., Angermeier, K., Harasymowyz, U., Schooley, K., & Belfiore, P. J. (2007). Effects of synthetic speech output on requesting and natural speech production in children with autism: A preliminary study. *Research in Autism Spectrum Disorders, 1,* 139–163.

Schlosser, R. W., & Wendt, O. (2008). Effects of augmentative and alternative communication intervention on speech production in children with autism: A systematic review. *American Journal of Speech-Language Pathology, 17,* 212–230.

Schmidt, J. D., Drasgow, E., Halle, J. W., Martin, C. A., & Bliss, S. A. (2014). Discrete-trial functional analysis and functional communication training with three individuals with autism and severe problem behavior. *Journal of Positive Behavior Interventions, 16,* 44–55.

Shane, H. C., Laubscher, E. H., Schlosser, R. W., Flynn, S., Sorce, J. F., & Abramson, J. (2012). Applying technology to visually support language and communication in individuals with autism spectrum disorders. *Journal of autism and developmental disorders, 42,* 1228–1235.

*Son, S. H., Sigafoos, J., O'Reilly, M., & Lancioni, G. E. (2006). Comparing two types of augmentative and alternative communication systems for children with autism. *Pediatric Rehabilitation, 9,* 389–395.

**Soto, G., Belfiore, P. J., Schlosser, R. W., & Haynes, C. (1993). Teaching specific requests: A comparative analysis on skill acquisition and preference using two augmentative and alternative communication aids. *Education and Training in Mental Retardation, 28,* 169–178.

*Tincani, M. (2004). Comparing the picture exchange communication system and sign language training for children with autism. *Focus on Autism and Other Developmental Disabilities, 19,* 152–163.

**Trembath, D., Balandin, S., Togher, L., & Stancliffe, R. J. (2009). Peer-mediated teaching and augmentative and alternative communication for preschool-aged children with autism. *Journal of Intellectual and Developmental Disability, 34,* 173–186.

*van der Meer, L., Didden, R., Sutherland, D., O'Reilly, M. F., Lancioni, G. E., & Sigafoos, J. (2012). Comparing three augmentative and alternative communication modes for children with developmental disabilities. *Journal of Developmental and Physical Disabilities, 24,* 451–468.

*van der Meer, L., Kagohara, D., Achmadi, D., O'Reilly, M. F., Lancioni, G. E., Sutherland, D., & Sigafoos, J. (2012). Speech-generating devices versus manual signing for children with developmental disabilities. *Research in Developmental Disabilities, 33,* 1658–1669.

*van der Meer, L., Kagohara, D., Roche, L., Sutherland, D., Balandin, S., Green, V. A., Sigafoos, J. (2013). Teaching multi-step requesting and social communication to two children with autism spectrum disorders with three AAC options. *Augmentative and Alternative Communication, 29,* 222–234.

van der Meer, L. A., & Rispoli, M. (2010). Communication interventions involving speech-generating devices for children with autism: A review of the literature. *Developmental Neurorehabilitation, 13,* 294–306.

*van der Meer, L., Sutherland, D., O'Reilly, M. F., Lancioni, G. E., & Sigafoos, J. (2012). A further comparison of manual signing, picture exchange,

and speech-generating devices as communication modes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 6,* 1247–1257.

Vannest, K., Parker, R. I., Gonen, O., & Adiguzel, T. (2016). Single Case Research: web based calculators for SCR analysis (Version 2.0) [Web-based application]. Texas A & M University. Retrieved from http://www.singlecaseresearch.org/

Vento-Wilson, M. (2014). Augmentative and alternative communication. In J. A. Ostergen (Eds.) *Speech-language pathology assistants: A resource manual* (pp. 381–399). San Diego, CA: Plural Publishing Inc.

What Works Clearinghouse. (2017). *Procedures and standards handbook* (Version 4.0). Retrieved from https://ies.ed.gov/ncee/wwc/

Received: 13 February 2019
Initial Acceptance: 9 April 2019
Final Acceptance: 15 May 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 142–157

Efficacy of a Computer-Based Editing Strategy with Postsecondary Students with Intellectual and Developmental Disabilities

Suzanne Woods-Groves
Auburn University

Charles A. Hughes
Penn State

Derek B. Rodgers
University of Nebraska-Lincoln

Kinga Balint-Langel
University of Minnesota Duluth

Saeed S. Alqahtani
Prince Sattam Bin Abdulaziz University

Katelyn M. Neil and Michelle Hinzman
University of Iowa

Abstract: The authors in this study used a pre-posttest experimental design with random assignment to treatment or control group to assess the use of an electronic editing cognitive strategy. The participants were 16 college students with intellectual and developmental disabilities enrolled in a 2-year postsecondary program at a Midwestern institute of higher education. Students who were taught the strategy received eight 50-minute lessons once a week for eight weeks. Each strategy lesson was driven by the strategy mnemonic and incorporated an explicit instruction format with modeling, guided practice with feedback, and independent practice. Strategy instruction included students using desktop PCs and Microsoft Word to identify and correct editing errors in electronic passages. Posttests revealed a significant difference in favor of the treatment group for total editing errors corrected and specific error types corrected for spelling, punctuation, and substance. No significant difference was found for overall appearance and capitalization errors. Two weeks following posttest, a maintenance probe revealed that students in the treatment group corrected a significantly higher number of editing errors than those in the control group.

Over the last decade there has been an upsurge in the number of postsecondary programs that support college students with intellectual and developmental disabilities (IDD) at institutes of higher education (IHEs) within the United States. Federal government incentives, such as the Higher Education Opportunity Act (HEOA) 2008 have afforded individuals with IDD improved access to higher education. It has been a decade since the enactment of the HEOA. Today potential college applicants with IDD can apply for federal financial aid along with their college-age peers without disabilities, and students with IDD have a greater number of IHE programs from which to choose (Grigal et al., 2018, 2017; Smith Lee, 2009). As the number of IHE programs that support students with IDD increase there is a pressing need to employ effective instructional strategies within these settings. Specifically, there is a need to identify strategies that can empower students with IDD to successfully engage in academic tasks such as writing in digital environments.

Writing in the 21st century requires one to have mastery of digital competencies when

The research reported herein was supported in part by the Office of Postsecondary Education (OPE), U.S. Department of Education, through Grant P407A100030 to The University of Iowa. The opinions expressed are those of the authors and do not represent views of the OPE or the U.S. Department of Education. A portion of work was conducted at the first author's previous affiliation, University of Iowa.

Correspondence concerning this article should be addressed to Suzanne Woods-Groves, Department of Special Education, Rehabilitation, and Counseling, 2084 Haley Center, Auburn University, Auburn, Alabama 36849. E-mail woodssu@auburn.edu

constructing and revising written text (DeVoss et al., 2010; Hayes & Olinghouse, 2015; Karchmer-Klein, 2013; Lewis, 2000; Regan et al. 2018). The 2010 *Common Core State Standards for English Language Arts & Literacy* support the use of technology in the writing process and note that college-ready writers must utilize technology in a strategic fashion when planning, constructing, and editing their own writing (National Governors Association Center for Best Practices & Council of Chief School Officers, 2010). Individuals with disabilities can benefit from strategy instruction that supports the use of word processing via computers or tablets to construct and edit written text (Englert et al., 2010; Graham et al., 2016; Graham & Perin, 2007; Silió & Barbetta, 2010). Current editing functions in word processing programs, such as spell check require the user to expand beyond accepting the provided answers and to engage in actions such as: (a) further examining the misspelled word if no acceptable answers are provided, (b) inspecting words that are homophones, and (c) reading the text aloud to identify misused words that are spelled correctly.

Postsecondary academic demands in writing span across a myriad of content and subject matter and require students to construct written assignments in an electronic form. Individuals with IDD who begin their IHE postsecondary experience with long standing difficulties in the area of written expression will continue to experience problems in academic writing tasks unless instructional support is provided. The writing process is a complicated iterative task that encompasses aspects of motivation, pre-planning, constructing a written product, and revision (Hayes, 2012; Hayes & Flower, 1987).

Individuals with IDD who struggle in written expression can experience problems with writing for a specific purpose, constructing organized written passages, revising their work, and difficulty in employing meta-cognitive skills (e.g. accessing and using writing strategies, evaluating one's own writing, revising and self-checking; Cannella-Malone et al., 2015; Connelly & Dockrell, 2016; Jackson et al., 2018; Joseph & Konrad, 2009; Konrad et al., 2006; Pennington, 2016; Pennington & Delano, 2012). As Hua et al. (2018) noted, young adults with IDD "may not derive the full benefits from postsecondary education" because of their persistent difficulties in literacy skills, such as constructing and revising written text (p. 1). In addition, postsecondary learners with IDD can experience barriers to effectively writing and revising even in electronic platforms, such as using computers to construct text due to lack of knowledge in how to use existing writing tools (e.g., spell checker, etc.). College students with IDD can benefit from using cognitive strategies that target editing and revising tasks in their electronic construction of texts which include identifying and correcting grammatical errors and making substantive corrections (Woods-Groves et al., 2017, 2015). These strategies provide students with mnemonic cues to aid in the recall of strategy steps and provide students with guidance in revising during the writing process (Hughes et al., 2010).

When writing within a digital platform such as using computers or tablets, previous studies have supported the use of strategy instruction in supporting learners with disabilities in utilizing digital tools included in most writing programs such as spell check, dictionary, thesaurus, and editing tools (Cullen et al., 2008; Graham et al., 2016; Graham & Perin, 2007; Hetzroni & Shrieber, 2004; Montgomery et al., 2001; Pennington, 2016; Wong, 2001). A meta-analysis conducted by Graham et al. (2016) revealed support for the following writing strategies for elementary and secondary learners: (a) teaching strategies to support word processing (d =1.46), (b) setting writing goals (d = .80), and (c) using explicit writing instruction in writing strategies (d = 1.26).

Cognitive strategy instruction is effective for students with and without disabilities in significantly improving writing and reading skills (Archer & Hughes, 2010; Conderman et al., 2013; Deshler & Schumaker, 1986; Englert; 1992; Graham, 2006; Harris et al., 2010; Hughes et al., 2017; Schumaker & Deshler, 2009). Cognitive strategy instruction in writing includes setting goals, selecting and successfully employing a learning strategy, monitoring oneself throughout strategy use, maintaining one's motivation throughout strategy use, and evaluating the execution of strategies (Arthur & Graham, 2016; Philippakos et al., 2015; Zimmerman, 2000). While empirical support for the efficacy of

strategy writing instruction with elementary and secondary learners with disabilities can be found, there is a need to investigate the use of cognitive strategy instruction with postsecondary learners with IDD in IHE settings.

A cognitive strategy instruction designed to teach editing skills within electronic platforms is the EDIT Strategy. The EDIT Strategy was developed by Hughes et al. (2010) through a combination of the Error Monitoring Strategy (Schumaker et al., 1985) and the InSPECT Strategy (Naughton & Hughes, 1999). The EDIT Strategy is a hybrid intervention that incorporates teacher-led explicit instruction, guided practice, and independent practice to teach students how to effectively use computer-based editing functions within word documents. The strategy was designed to teach students how to identify and correct grammatical errors in their digital word documents. EDIT is a mnemonic device which represents the four steps of the strategy: a) Enter your first draft; b) Do a spell check; c) Interrogate yourself using the capitalization, overall appearance, punctuation, and spelling (COPS) questions, and d) Type in corrections and run the spell-checker (Hughes et al., 2010). In addition to the mnemonic, the EDIT strategy uses explicit instruction procedures (Archer & Hughes, 2009), mastery-based learning (Schumaker & Deshler, 2009), and goal setting and self-evaluation (Schunk & Zimmerman, 2007) to guide participants through the digital editing process.

There have been three previous investigations of the EDIT intervention to date. Carranza and Hughes (2009) first studied the effects of EDIT with 22 elementary and middle school students with learning disabilities (LD) who were randomly assigned to treatment and control groups. The study lasted three weeks with a total of four instructional hours. The lessons included modeling of the intervention with teacher think-alouds, practice of the four EDIT steps, teacher guided practice with feedback and error correction, and independent practice. The dependent measure was a pre- and posttest electronic passage that contained 25 total errors, five for each of the EDIT categories: Spelling, Capitalization, Overall Appearance, Punctuation, and Substance. Treatment and control groups were compared on their EDIT Total score (i.e., the total number of errors corrected) and on their performance for each of the five error categories. A comparison of the posttest probes revealed that the treatment group outperformed the control group. In addition, the treatment group improved upon their pre-test scores from 28% to 80% of total errors corrected, and they maintained their skills when assessed several weeks after the study concluded.

Woods-Groves et al. (2015) studied the effectiveness of the EDIT strategy with 19 postsecondary students with intellectual and developmental disabilities (IDD). All participants were enrolled in a campus-based, 2-year certificate program at an Institute of Higher Education located in the Midwest. The investigators used a pre- and posttest design, randomly assigning participants to a treatment ($n = 11$) or control ($n = 10$) group. The strategy was taught in sixteen 50-minute lessons over eight weeks, for a total instructional time of 13.33 hours.

The dependent measure consisted of two electronic Microsoft word passages (i.e., California Redwoods and Giant Pandas) each from the EDIT Strategy manual and materials (Hughes et al. 2010). The two passages were adapted to the third-grade reading level and contained approximately 220 words each. Each passage contained 25 errors with five errors in each of the following areas: spelling, capitalization, overall appearance, punctuation, and substance. Scores from the prompts included a total score and scores for the five error types. Students were randomly assigned one of the two passages as a pretest and were assigned the remaining passage as a posttest. For instance, if one student was randomly assigned California Redwoods for a pretest then the student would have Giant Pandas as a posttest. This served as a counterbalance for the pre- and posttest assignments across participants. The treatment group performed significantly better than the control group on three areas: EDIT Total Score ($d = 1.01$, $p = .011$), Overall Appearance ($d = 1.06$, $p = .048$), and Punctuation ($d = 1.54$, $p = .004$). In addition, the treatment group maintained significant differences in performance on two areas 11 weeks after the intervention had con-

cluded: EDIT Total (d = 1.19, p = .029) and Overall Appearance (d = 1.67, p = .004).

Woods-Groves et al. (2017) investigated the EDIT strategy with 15 postsecondary students with IDD. Woods-Groves et al. (2017) shortened the overall instruction time reported in Woods-Groves et al. (2015) by reducing lessons pertaining to using the spell checker. In this investigation the participants were randomly assigned to treatment (n = 7) and control (n = 8) groups with a pre-/posttest design. The EDIT strategy was delivered to the treatment group in 11 sessions that lasted 45-minutes over 5.5 weeks. The total instructional time was 8.25 hours. Woods-Groves and colleagues extended the pre- and posttest probe ranges by adding five more grammatical errors to the prompts for each error category which increased the total errors from 25 to 30 per probe. All prompts were written at the third-grade level and averaged 220 words each.

Even with the shorter instructional time, the researchers found significant differences between the treatment and control groups on four measures: EDIT Total (d = 0.84, p = .006), Spelling (d = 0.71, p = .022), Punctuation (d = 0.96, p = .030), and Substance (d = 1.66, p = .007). The investigation also included two maintenance periods, one at five weeks after the intervention and another at 12 weeks post-treatment. The five-week maintenance phase revealed significant differences in measures identical to the post-test: EDIT Total (d = 1.01, p = .003), Spelling (d = 0.97, p = .009), Punctuation (d = 1.17, p = .023), and Substance (d = 1.13, p = .022). The 12-week maintenance phase comparisons yielded significant differences in three areas: EDIT Total (d = 0.41, p = .014), and Spelling (d = 0.60, p = .042), Punctuation (d = 0.90, p = .025).

In the previous two EDIT investigations conducted by Woods-Groves and colleagues the findings supported the use of the strategy with college students with IDD, however the results were not conclusive. In Woods-Groves et al. (2015), students who were taught the strategy (treatment group) significantly corrected a higher number of overall appearance and punctuation errors than students in the control group. In the second EDIT investigation with college students with IDD, students who were taught the strategy significantly outperformed those who were not taught the strategy in the number of corrected spelling and punctuation errors. Neither study revealed an improvement in the area of substantive errors corrected while only the area of punctuation showed consistent effects across the two studies.

It is important to identify potentially effective strategies to improve the writing skills of students with IDD enrolled in postsecondary programs. Potential strategies must be experimentally tested within authentic settings with the target population for which they are intended to be used. This current investigation is designed to further examine the use of the EDIT strategy with college students with IDD.

The following research questions were investigated:

1. Will participants in the EDIT Strategy instructional group independently correct a higher number of editing errors in electronic word passages than participants in the control group?
2. Will participants in the EDIT Strategy instructional group independently correct a higher number of editing error types including spelling, capitalization, overall appearance, punctuation, and substance in electronic word passages than participants in the control group?
3. Will participants who received the EDIT Strategy maintain their posttest performance during the maintenance stage?

Method

Participants

Sixteen postsecondary students with IDD participated in this study. A convenience sample was used in this study where all participants were young adults enrolled in a 2-year postsecondary campus-based program for individuals with IDD at a research one Midwestern university. The participants ranged in age from 20-to-22 years (M = 18.88, SD = 1.15). With regard to educational services, the participants were served under the following diagnostic categories: 2 (12.5%) with Autism, 2 (12.5%) with Pervasive Developmental Disorder-Not Otherwise Specified, 1 (6.3%) with

Fetal Alcohol Syndrome and Attention Deficit Disorder, 8 (50%) with an intellectual disability, 2 (12.5%) with an intellectual disability and cerebral palsy, and 1 (6.2%) with a learning disability and cerebral palsy.

There were seven (44%) females and nine (56%) males. With regard to race, there were 15 (94%) participants who were White and one (6%) participant who was Asian. The students were from rural 8 (50%), suburban 4 (25%), and 4 (25%) urban areas. All participants were assessed with the Woodcock Johnson Tests of Achievement III (WJIII; Woodcock et al., 2001). Across all the participants, WJIII Total Scores, with a mean of 100 and standard deviation of 15, ranged from 39 to 104, (*Mdn,* 72), while WJII Broad Reading scores ranged from 34 to 115, (*Mdn,* 53), and WJIII Broad Writing scores ranged from 28 to 116, (*Mdn,* 60). The WJII Total Scores were used to stratify all the participants, then a coin flip (i.e., tails = treatment group; heads = control group) was employed to randomly assign participants to treatment or control groups. There were eight students respectively assigned to the treatment and control groups.

Material

Materials from Hughes' et al. (2010) EDIT Strategy manual were used for this study. Six adaptations were made to the original EDIT Strategy materials. First, all the EDIT Strategy prompts for pre- and posttests and maintenance, and prompts used in lessons were adapted to the lowest reading level indicated across the participants. The participants' curriculum-based measurement oral reading fluency results via the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) oral reading fluency (i.e. DORF) indicated that the lowest reading level for all students was the third-grade level range. As a result, all EDIT Strategy materials were adapted to not exceed the third-grade level as determined by Flesh-Kincaid readability tests (Kincaid et al., 1975). Second, materials employed in each intervention lesson included an EDIT Strategy graphic organizer that contained the steps of the respective lesson for the session and a highlighter that students could use to check off each completed step on their graphic organizer. Third, USB memory sticks with electronic word files (i.e., guided practice, independent practice passages) for the respective lesson were provided to each student. Fourth, each student was given a folder with a graph to self-graph and monitor performance. The fifth adaptation to the strategy included providing the students with headphones and instructed to use the text-to-speech universal design function when reviewing their edit prompt during the strategy. The sixth and final adaptation to the strategy was the extension of the original four-hour total instructional time noted in Carranza and Hughes' (2009) investigation with students with LD.

During the EDIT Strategy intervention, the students were instructed in a large group format in a computer lab located at the university. This computer lab was a classroom that routinely held college classes for students enrolled in the university. The computer lab had four long tables with approximately five computers stationed at each table. The tables were arranged where students faced the front of the classroom where a smartboard, instructor lectern, document camera, and an overhead projector were located. Each student had a PC computer with a mouse, USB ports, and a monitor. The instructor was employed through the postsecondary program and regularly taught all the students in other courses within the postsecondary program (i.e., career development, finance, etc.). For each EDIT Strategy lesson the instructor employed a projector, Smartboard, and a PC desktop computer at the front of the computer lab to teach each lesson. See EDIT Graphic Organizer in Figure 1.

Design and Procedure

Design. An experimental pre-/posttest design was employed in this study. A coin flip procedure was conducted to assign 16 students to treatment or control groups. As noted previously, the students were stratified based upon their WJIII Total Scores then assigned via coin flip to respective groups (i.e., tails = treatment group; heads = control group). There were eight students in each group (i.e., treatment and control). The WJIII Total Scores for the participants that were assigned to respective groups were compared via an Analysis of Variance (ANOVA) and re-

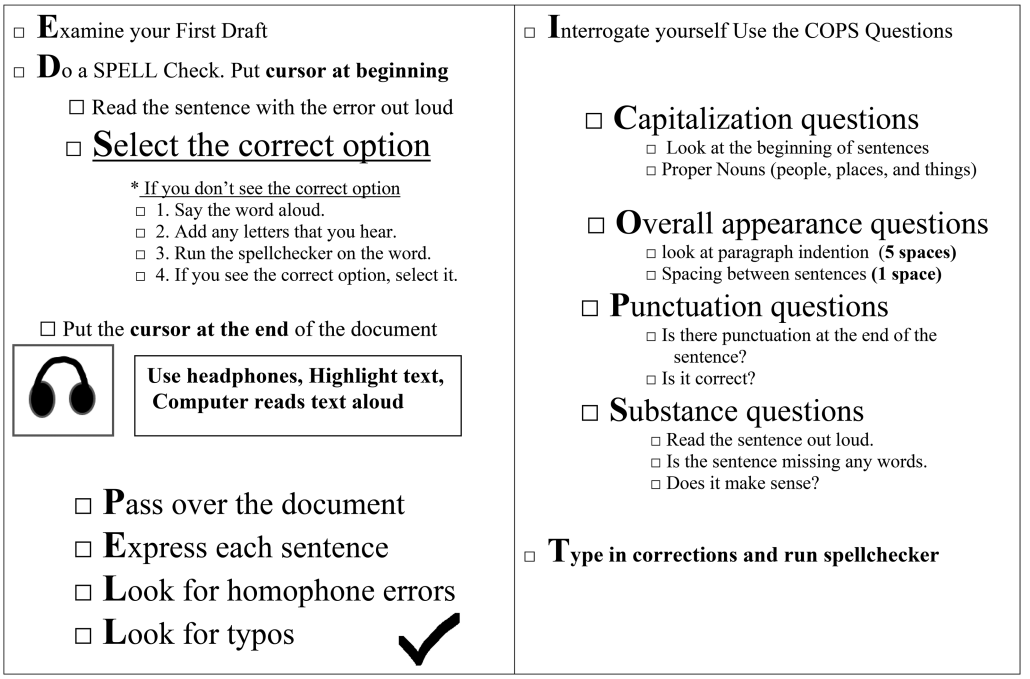


Figure 1. EDIT Strategy Graphic Organizer (adapted from Hughes et. al., 2010).

vealed no significant differences between the groups $F(1, 15) = .160$, $p = .695$, treatment group ($M = 68.38$), control group ($M = 72.50$).

The participants in the treatment and control groups were administered a pretest one week preceding the beginning of the editing intervention and a posttest one week following the end of the intervention. The pre- and posttests were the "California Redwood" and the "Giant Panda" prompts. These two prompts were randomly assigned to participants in treatment/control groups. Administration of the prompts was counterbalanced. For example, if the "California Redwood" prompt was assigned to an individual as a pretest, then that individual was given the "Giant Panda" prompt for their posttest. The participants were administered a maintenance prompt three weeks following instruction and two weeks following the posttest.

The pre- and posttest prompts were adapted electronic versions of the "California Redwood" and "Giant Panda" prompts provided in the EDIT Strategy manual (Hughes et al., 2010). The prompts were revised to not exceed a third-grade readability level range as noted earlier. Each prompt consisted of a passage with approximately 220 words each. And contained 30 errors with six errors in each of the following areas: spelling, capitalization, overall appearance, punctuation, and substance. For each prompt the passage consisted of six spelling errors including two homophone and four words dispersed within the passage that were misspelled. Capitalization errors included a total of six errors for the beginning of the sentence and or proper nouns within the passage. Overall appearance errors included six errors consisting of errors in indention for paragraphs and errors for spacing between sentences. Punctuation errors included six errors throughout the passage that included the use of punctuations, such as periods and question marks. Substance errors were dispersed throughout the passage and included six instances of missing articles (e.g. the, a, and) and subjects. Each passage was scored using the EDIT Scoring key designed for each prompt that provided an answer key for the five error types for a total of 30 errors. In addition, each EDIT

Errors	Spelling ✓		Capitalization	Overall Appearance	Punctuation	Substance Missing Word
25	5	1 Homophone	5	4	5	5
Errors Corrected						
Errors Not Corrected						

_____ correct out of 25 = _____% _____total errors

Circle errors in document

Figure 2. Maintenance EDIT Strategy Scoring Key.

Scoring key provided a way to score and record each error type, and a rubric for totaling correct and incorrect responses. The maintenance prompt EDIT Scoring key was similar to the California Redwoods and Giant Panda EDIT Scoring keys, except that the maintenance scoring key had consisted of 25 errors with five errors for each error type. See Figure 2 for the maintenance EDIT Scoring key.

All assessments (i.e., pre-/posttest prompts and a maintenance prompt) were administered in a large group format in a computer lab. Each student had the respective electronic assessment on their PC desktop in a Microsoft Word document. Instructions included the following: "Please open the word document that has been placed on your desktop. You have 30 minutes to read the prompt and then run the spell checker. You are to correct any errors you find in your prompt" (adapted from Hughes et al., 2010). No participants performed above the study exclusion criterion of 80% or more total errors corrected in the pretest and were subsequently all included in the study.

Intervention. Several adaptations in the EDIT Strategy were implemented in this study that diverge from previous published investigations of the EDIT Strategy with college students with IDD (Woods-Groves et al., 2017, 2015). Adaptations included the following: (a) a reduction of the number of lessons taught in previous investigations of the EDIT Strategy with young adults with IDD from 16 and 11 lessons to eight 50 min. lessons once a week, (b) the use of assistive technology/universal design features in Microsoft Word (i.e., text to speech and headphones), and (c) the use of adapted graphic organizers to support strategy steps. These adaptations were designed to decrease the length of instructional time and to support students in their use of the strategy.

EDIT Strategy lessons were conducted once a week for eight 50-minute sessions. Total intervention time was 6 hours and 40 minutes. The instructor for the EDIT Strategy was a graduate student in counseling and was employed by the postsecondary IDD program as an official teacher in the program. The first author met weekly with the EDIT Strategy instructor to discuss aspects of upcoming lessons and instruction.

The sequence of each of the eight EDIT Strategy lessons is described below. Specific lesson content is included in Table 1. Aspects encompassed in each lesson included the following: (a) explicit instruction with modeling, guided practice with feedback and independent practice, (b) mastery-based learning to an 80% criterion before advancing to a new strategy step, (c) respective graphic organizers that addressed steps previously taught and current lesson steps, (d) highlighters, and (e) student folders with the EDIT mnemonic on the front, self-graphing table, and a guide for how to access and save electronic documents within each folder. Students were each given a point booklet that included the class expectations, such as arriving on time, completing independent practice with 80% mastery, and participating in class. Students graphed their performance for independent practice and critiqued their behavior regarding following class expectations by assigning themselves points within their point booklet at the end of each session. Points for the class were cumulatively summed each week and "banked" for the goal of having a pizza party.

TABLE 1

Components of EDIT Strategy Lessons (adapted from Hughes et al., 2010)

Lesson	*Sessions*	*Lesson Activities*
1	1	The instructor defined editing and asked students to identify difficulties they encounter when writing and editing. Students discussed strategies they use to edit their work, including the use of spell checker in Microsoft Word. The instructor introduced the EDIT mnemonic and strategy and explained the benefits of using the strategy.
2	3	The instructor introduced the first two steps of the EDIT strategy (i.e., "Examine your first draft," and "Do a SPELL Check"). Students learned how to use spellcheck to identify errors in a Microsoft Word document and used the following procedure to correct errors. First, check to see if spellcheck suggested the correct word. If not, sound out the incorrectly spelled word, change the spelling, and run the spellchecker again. If the correct spelling is still not available, use another word that has the same meaning as the incorrectly spelled word. Students were instructed to use the headphones and "speak" function Microsoft Word during the SPELL steps to have each sentence read aloud beginning at the end of the passage and moving through sentences from the end of the text to the beginning of the text. As students listen to each sentence they are to complete the rest of the SPELL steps (i.e. pass over the document, express each sentence, look for homophone errors, and look for typos). Students learned to save their work on a USB memory stick.
3	1	Students reviewed the previous two steps of the EDIT strategy (i.e., Examine your first draft, Do a SPELL Check) and graphed their graded work from the previous lesson. The instructor introduced common types of writing errors using the COPS mnemonic (i.e., Capitalization errors, Overall appearance errors, Punctuation errors, or Substance errors). The instructor used think-alouds to model the COPS steps, and students practiced with instructor guidance. At the end of the lesson, the instructor introduced the last step of the EDIT strategy (i.e., Correcting typos and running the spell checker once more). Students practiced saving their work on their USB memory sticks.
4	3	Students reviewed all steps of the EDIT strategy, including the SPELL and COPS mnemonics. Students then practiced using the EDIT strategy on electronic passages under the guidance of the instructor. The instructor provided each student individually with corrective feedback immediately after they completed the passage. Students then completed passages independently.

The type of editing errors addressed in the EDIT Strategy instruction included: spelling (including homophones), capitalization, overall appearance, punctuation, and substantive errors. For this investigation the EDIT Strategy was adapted by adding the use of the universal design features in Microsoft Word, such as the text to speech (i.e. "speak" function) and the use of headphones. An adapted graphic organizer was used in each lesson that depicted the strategy steps.

Two staff members of the postsecondary program were trained to collect treatment integrity data for each session. The first author provided a 30 min training to the staff members. The first author explained each step of the lesson and showed the staff members where each step was aligned with the respective item on the treatment integrity rating sheet. Mastery of training was determined by the first author based on the staff members' correct responses (e.g. all responses 100% correct) to the first author's questions concerning the lesson scripts and corresponding treatment integrity rating components. For instance, for the treatment integrity item "instructor provided advance organizer" the first author would ask the staff members to "describe an example from the script of what the instructor would say in order to provide the advance organizer." During the EDIT Strategy intervention sessions, the staff members were provided a copy of the instructor script and a list of consecutive steps that were to be com-

pleted for each respective lesson. Two staff members collected treatment integrity for six of the eight EDIT Strategy sessions. For the first and seventh session one rater collected treatment integrity due to rater absence, and the treatment integrity checklists indicated 100% fidelity for each session. For the remaining six sessions where treatment integrity was collected via both raters, both raters had 100% agreement and indicated 100% fidelity for each session.

Control group intervention. Participants in the control group who did not receive EDIT Strategy instruction participated in Science instruction in a large group format at the same time EDIT Strategy instruction was delivered. All participants regularly attended classes delineated in their postsecondary program throughout the day, five days a week. These classes pertained to transition and academic instruction (e.g., finance, career instruction).

Dependent variable EDIT Strategy scoring keys. Pre-, posttest, and maintenance prompts were all scored using the EDIT Strategy scoring keys. With regard to the edit prompts, Hughes et al. (2010) pre- and posttest prompts were adapted from the original EDIT Strategy testing materials provided in the manual. This investigation employed pre- and posttest prompts that had been adapted to not exceed the third-grade range in reading and that each included 30 total editing errors (i.e., six respective examples of spelling, capitalization, overall appearance, punctuation, and substance). The two adapted pretest prompts pertained to two different topics. The first prompt was entitled "Giant Panda" and pertained to characteristics of Giant Pandas. The second prompt was entitled "California Redwood" and pertained to facts about California Redwoods.

The pre- and posttests and maintenance prompt were all scored using respective EDIT Strategy scoring keys. The "California Redwood" and "Giant Panda" EDIT Strategy scoring keys each contained answers for the 30 editing errors for each respective prompt. The possible Total Score ranged from 0 to 30. There were six editing error examples for each edit error type (i.e., spelling, capitalization, overall appearance, punctuation, and substance). Possible scores for each editing type ranged from 0-to-6. As noted earlier, specific editing error types for spelling included two homophones and four incorrectly spelled words. Capitalization errors pertained to the beginning of a sentence or proper nouns. Overall appearance errors included additional spacing or omitted spacing between sentences and paragraphs. Substance errors pertained to missing subjects and articles.

A maintenance prompt was created by following the format of the EDIT Strategy manual original prompts. The maintenance prompt was written at the 2.3 grade level using Flesh-Kincaid and included 25 errors (i.e., six spelling errors, five capitalization, four overall appearance, five punctuation, and five substance). The prompt was a narrative passage about an adolescent named "Jake" who compared the advantages of driving different types of passenger vehicles. The maintenance prompt EDIT Strategy scoring key contained correct answers for the 25 errors with a possible Total Score that ranged from 0-to-25, and a respective range of scores (i.e. spelling 0-to-6, capitalization 0-to-5, overall appearance 0-to-4, punctuation 0-to-5, and substance 0-to-5).

Data collection. One week before EDIT Strategy instruction began all participants were administered a pretest (i.e., counter-balanced "California Redwood" and "Giant Panda" prompts) in a large group setting, in a computer lab. One week following the end of EDIT Strategy instruction, a posttest was administered (i.e., counter-balanced "California Redwood" and "Giant Panda" prompts) to all participants. The participants in the treatment and control groups completed a maintenance test two weeks following completion of the posttest.

Two graduate students enrolled in the College of Education at the Midwestern University where the study was conducted were trained by the first author in completing EDIT Strategy scoring keys. The one-hour training consisted of the first author providing the raters with examples of "fabricated samples" of completed edit prompts and the respective scoring keys. The first author reviewed the samples with the raters. Each rater scored the completed samples using the EDIT Scoring Keys. Raters' fidelity of scoring was achieved when the raters' scores were 100% correct and

when there was a 100% agreement between the raters.

Once the raters were trained they proceeded to score all completed prompts. The raters were blind with regard to whether completed prompts they were scoring were completed by treatment or control participants. The raters also did not know if the prompts were pre-, posttest, or maintenance prompts. All completed prompts were scored by these two raters who had previous experience in administering and evaluating assessment materials.

Data Analysis

Previous experimental group studies pertaining to the EDIT Strategy yielded Cohen's *d* effect sizes for EDIT Total Errors corrected as .84 and 1.01 (Woods-Groves et al., 2017; Woods-Groves et al., 2015). A power analysis was conducted using G power (Faul et al., 2007) with an effect size of .80 (based on previous EDIT studies' results) and an alpha of .05. G power results indicated a minimal sample size for two groups of 15 would be adequate ($N\_14.6429 - N\_15$). The results from the raters' completed respective Scoring Keys for the pretest and posttest prompts and the maintenance prompt were analyzed via SPSS 23 (2016). Bivariate correlations between raters' results were conducted. A series of ANOVA and analysis of covariance (ANCOVAs) with the pretest as a covariate were conducted with the average of the raters' results for the Scoring keys for pretests, posttests, and maintenance. Effect sizes were calculated via Cohen's *d* for all analyses. Cohen's (1988) criteria were applied. Effect sizes of $<.2$ were determined to be small, while results of $<.5$ were considered medium, and results of $>.8$ were considered large.

Results

Treatment Integrity and Inter-Rater Reliability

Treatment integrity checklists were completed for each of the eight EDIT Strategy lessons. There was 100% compliance with all the steps being followed for each of the eight lessons conducted. The mean from the two raters was calculated for each aspect of the EDIT Strategy scoring keys (i.e., pre-/posttests and maintenance). Inter-rater reliability for the raters' scores were compared via bivariate correlations for the pre-/posttests, and maintenance. The range and median Pearson product moment correlations for the Total EDIT score and each of the five editing error types included the following: (a).82 to .99 (Mdn = .96) for the pretests; (b) .89 to .99 (Mdn = .96) for the posttests; and (c) .99 to 1.00 (Mdn = .94) for the maintenance.

EDIT Strategy scoring key. The EDIT Strategy scoring key results for pretests for the control and treatment groups were examined via a series of ANOVAs. The average of the raters' scores for the EDIT Strategy scoring keys revealed no significant results when the treatment and control groups were compared for the pretest Total EDIT Errors corrected and five error types. The posttest EDIT Strategy scoring key results for posttests for the treatment and control groups were compared. The average of the raters' EDIT Strategy scoring keys for treatment and control groups was examined via ANCOVAs with the pretests as a covariate. The posttest Total EDIT Errors corrected revealed a significant positive difference for the treatment group when compared to the control group $F(2, 15) = 15.175$, $p = .002$ with a large effect size $d = 1.38$. An examination of the posttest specific editing error types revealed significant positive differences for spelling, punctuation, and substance for the treatment group when compared to the control group. No significant results were found for the error types overall appearance and capitalization. Pre- and posttest results are depicted in Table 2.

Maintenance. Two weeks following the posttest the students were given an EDIT Strategy electronic passage in the same format as the pre- and posttest passages. The setting was two computer labs at the university. Students in the treatment group completed the passage in a separate computer lab than the students in the control group. All the students were provided with a desktop PC and Microsoft Word to access the passage. The instructor provided a brief review of the EDIT Strategy for the students who were taught the strategy. Both groups were instructed to find and correct all editing errors in the passage. Two raters previously trained in how to score

TABLE 2

EDIT Strategy Assessments Pre- and Posttests Results

Pre/Post Test	*Total Errors Corrected*	*Spelling*	*Capitalization*	*Overall App*	*Punctuation*	*Substance*
C group (8)	17.88* (4.26)	4.88* (1.25)	5.13* (0.99)	4.56* (1.18)	2.63* (1.99)	0.50* (0.71)
T group (8)	16.31* (2.45)	5.00* (0.71)	5.31* (0.88)	4.81* (1.25)	1.06* (1.05)	0.13* (0.35)
Cohen's d	.45	.12	.19	.21	.99	.66
ANOVA	$F(1,15) = .810$	$F(1,15) = .061$	$F(1,15) = .159$	$F(1,15) = .169$	$F(1,15) = 3.841$	$F(1,15) = 1.800$
	$p = .383^{ns}$	$p = .809^{ns}$	$p = .696^{ns}$	$p = .687^{ns}$	$p = .070^{ns}$	$p = .201^{ns}$
Posttest						
C group (8)	16.94* (5.05)	4.75* (1.16)	5.06* (1.02)	4.00* (1.51)	2.25* (2.12)	0.81* (1.39)
T group (8)	23.81* (4.88)	5.88* (0.35)	4.38* (1.85)	4.88* (1.43)	5.25* (1.00)	3.19* (1.89)
Cohen's d	1.38	1.32	.46	.59	1.81	1.43
ANCOVA	$F(1,14) = 15.175$	$F(1,14) = 12.773$	$F(1,14) = .733$	$F(1,14) = 1.667$	$F(1,14) = 17.297$	$F(1,14) = 15.138$
	$p = .002^{**}$	$p = .003^{**}$	$p = .407^{ns}$	$p = .219^{ns}$	$p = .001^{**}$	$p = .002^{**}$

Note. * Denotes mean values, C group (8) = number of control group participants, T group (8) = number of treatment group participants, Standard deviations are provided in parentheses. Overall App = Overall Appearance. ^{ns} = non-significant. ** Denotes significant. ES = Effect size, Cohen's d.

passages using the EDIT Scoring keys scored the students' responses. The average of the two raters' Maintenance EDIT Scoring Key results were compared for the treatment and control groups via an ANCOVA with the pretest results as a covariate. Results revealed a significant difference in favor of the treatment group for the EDIT Total Score group $F(1, 14) = 5.215$, $p = .040$ with a large effect size $d = .93$, however, no significant results were found for specific error types when the treatment and control groups were compared. The maintenance results including means, standard deviations, ANCOVA, and effect sizes are depicted in Table 3.

Discussion

When college students with IDD enter postsecondary programs at IHEs they are investing their time and financial resources. It is essential that students with IDD who have existing difficulties in the writing process including editing and revising within digital environments are provided with effective instruction within the IHE setting. While there is evidence of the positive effect of cognitive strategy instruction in improving writing skills for elementary and secondary learners with disabilities, there is a need to investigate the efficacy of writing strategies with postsecondary learners with IDD.

The purpose of this investigation was to further examine the use of the EDIT Strategy with postsecondary learners enrolled in a two-year college program for individuals with IDD. The first research question we investigated pertained to: Will students who were taught the EDIT Strategy significantly improve their skill in identifying and correcting editing errors within electronic documents when compared to the control group? Results from this study supported the use of the EDIT Strategy with the students who were taught the strategy, significantly out-performing students in the control group in the number of Total Editing Errors corrected, yielding a large effect size, $d = 1.38$. This finding was comparable to the two previous experimental group studies conducted by Woods-Groves and colleagues with postsecondary students with IDD ($n = 34$, across both studies) where effect sizes for Total Editing Errors corrected ranged

TABLE 3

Maintenance Probe: Treatment/Control Groups for 2 Weeks Following EDIT Strategy Instruction

Administration	*Total Errors Corrected*	*Spelling*	*Capitalization*	*Overall App*	*Punctuation*	*Substance*
MP 2 Weeks						
C group (8)	15.63* (4.11)	5.63* (.52)	4.63* (0.52)	1.81* (1.13)	2.69* (2.05)	0.88* (1.13)
T group (8)	19.31* (3.76)	6.00* (0.00)	4.81* (0.53)	2.38* (1.09)	3.31* (1.33)	2.81* (2.14)
Cohen's *d*	.93	1.01	.34	.51	.36	1.13
ANCOVA	F (1,14) = 5.215	F (1,14)= 3.987	F (1,14) = .446	F (1,14) = .795	F (1,14) = 2.356	F (1,14) = 4.121
	p = .040**	p = .067^{ns}	p = .516^{ns}	p = .389^{ns}	p = .149^{ns}	p = .063^{ns}

Note. * Denotes mean values, C group (8) = number of control group participants, T group (8) = number of treatment group participants, Standard deviations are provided in parentheses. Overall App = Overall Appearance. ^{ns} = non-significant. ** Denotes significant. ES = Effect size, Cohen's *d*. MP = Maintenance Probe.

from d = .84 to 1.01 (Woods-Groves et al., 2017, 2015).

In the current investigation students received 6 hr. and 40 min of instruction. Positive effects were found despite the reduction in instructional time from the reported 13.3 hr. from Woods-Groves (2015) and 8.25 hr. from Woods-Groves (2017). Adaptations to the EDIT Strategy instruction included the use of text to speech and headphones during the revising process and eliminating one SPELL Check session and one COPS session from the reported number of sessions in Woods-Groves (2017). The reduction in instructional time is important to note given the need to identify effective yet usable and feasible instructional strategies within authentic IHE settings.

The second research question investigated was: Will students who were taught the EDIT Strategy correct a significantly higher number of error types (i.e., spelling, capitalization, overall appearance, punctuation, and substance) when compared to the control group? In this investigation we found a significant difference in the treatment groups' number of spelling, punctuation, and substance errors corrected when compared to the control group, but no significant differences in overall appearance and capitalization errors corrected. When these results were compared with the two previous postsecondary group EDIT Strategy studies, commonalities were revealed for some error types, while disparate results were found for others. In all three postsecondary studies students who were taught the strategy corrected a significantly higher number of punctuation errors with effect sizes that ranged from d = .96 to 1.81 with no significant effect found for capitalization errors. However, only one study (i.e., Woods-Groves, 2015) reported significant effects for overall appearance errors corrected. Two studies (i.e. the current investigation and Woods-Groves, 2017) revealed significant effects for spelling errors (d = .71 to 1.32) and substance errors (d = 1.41 to 1.66). It is important to note that while the instructional time was shortened in the current investigation, a significant difference and large effect size was found for the correction of substance errors, which is a more complicated revising skill. The third research question pertained to: If students who were taught the strategy

would retain their posttest skill level two weeks following the end of instruction? The maintenance probe administered two weeks following the posttest revealed that students in the treatment group corrected a significantly high number of total errors, but not respective error types when compared to the control group. Differences in the mean scores and effect sizes for the treatment group indicate an elevated number of error types were corrected when compared to the control group, however these differences were not significant. The two previous postsecondary IDD studies indicated that students who were taught the strategy significantly outperformed students in the control group in total errors corrected and in correcting error types in one of the two areas where they previously excelled 12 and 11 weeks respectively following posttest.

Limitations and Future Research

There were five limitations noted in this investigation. The first limitation pertains to the fact that we did not see a significant improvement in the correction of all editing error types when treatment and control groups were compared. In addition, students in the treatment group did not retain their posttest level of performance when tested two weeks later. This could be due to an instructional error where the number and focus of sessions may need to be increased or improved in areas (e.g., frequency of corrective feedback in guided practice). It could be due to a flaw in our dependent measure, the maintenance probe. While the pre- and posttest prompts each had 6 errors for each of the five error types, the maintenance probe only afforded students 25 error correction opportunities (i.e., six spelling errors, five capitalization, four overall appearance, five punctuation, and five substance). There may be a restriction in range with regard to the instances of error types in the maintenance probe used for this investigation. The second limitation was that students were not able to apply their editing skills to correcting their own work. Due to time constraints and the multi-faceted mnemonic steps taught in the strategy we were not able to have students practice their editing skills with their own written products. Future studies could perhaps embed an extension activity following independent practice that would have students generalize the editing skills they learn in each lesson to their own written work. A necessary next step in examining the EDIT Strategy is for the intervention to be used by individuals to edit their own work. The third limitation was that while we were able to examine the use of the EDIT Strategy in an authentic IHE setting, instructional settings, such as one-to-one tutoring, or small group instruction within an IHE inclusive classroom should be examined in future studies. The fourth limitation pertained to the use of a coin flip to assign participants to control or treatment groups. The use of a coin flip for assignment could result in an imbalance in participants with similar attributes being over assigned to one of the two groups (Kang et al., 2008). In order to address this limitation participants were stratified based on their standardized achievement scores (i.e., WJII Total Scores) before being assigned via coin flip. While there still could be an imbalance in participant assignment, stratification was one way to aid in the random assignment procedure (Kernan et al., 1999). The fifth and final limitation pertained to the fact that social validity data were not collected in this investigation concerning the students' perceptions of the EDIT Strategy, including the usability and feasibility of using the strategy in college settings. It is important to gather social validity information when conducting intervention studies in authentic settings to assess attitudes and views of students who use the strategy or methods. Social validity information can inform the further development of an intervention (Carroll & St. Peter, 2014). Future studies should include a metric to gather information from students and instructors concerning the use of the strategy.

Implications for Practice

Requiring college students with IDD who have difficulty in written expression to complete writing compositions with computers or tablets without further technological guidance is a disservice to the students. Not providing students with IDD support in accessing technology for writing can be construed as a

missed opportunity for teaching students to access universal design features of existing writing programs, such as spell check, text to speech, highlighting functions, and use of a dictionary or thesaurus. Cognitive strategy instruction that pertains to writing and revising in digital environments may afford postsecondary students with IDD an effective support to improve their performance in constructing and revising written text. To date, there have been four experimental investigations of the EDIT Strategy, one with upper elementary students with LD and three with postsecondary learners with IDD. Each study has supported the use of the EDIT Strategy in improving the editing skills of students who were taught the strategy. While empirical support is beginning to build, it is imperative that effective instructional practices be identified and examined for their use with young adults with IDD in authentic IHE settings.

References

Archer, A. L., & Hughes, C. A. (2010). *Explicit instruction: Effective and efficient teaching*. The Guilford Press.

Cannella-Malone, H., Konrad, M., & Pennington, R. C. (2015). ACCESS! Teaching writing skills to students with intellectual disability. *Teaching Exceptional Children, 45,* 272–280.

Carranza, M., & Hughes, C. A. (2009). *Effects of teaching an editing strategy to middle school students with learning disabilities*. Unpublished manuscript.

Carroll, R. A., & St. Peter C. C. (2014). Methods for assessing social validity of behavioral intervention plans for children with attention deficit hyperactivity disorder, *Acta de Investigación Psicológica, 4,* 1642–1491.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.

Conderman, G., Hedin, L., & Bresnahan, V. (2013). *Strategy instruction for middle and secondary students with mild disabilities: Creating independent learners*. Sage.

Connelly, V., & Dockrell, J. (2016). Writing development and instruction for students with learning disabilities. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research,* (2nd ed.; pp. 349–363). The Guilford Press.

Cullen, J., Richards, S. B., & Frank, C. L. (2008). Using software to enhance the writing skills of students with special needs. *Journal of Special Education Technology, 23*(2), 33–44.

Deshler, D. D., & Schumaker, J. B. (1986). Learning strategies: An instructional alternative for low-achieving adolescents. *Exceptional Children, 5*2, 583–590.

DeVoss, D. N., Eidman-Asdahl, E., & Hicks, T. (2010). *Because digital writing matters: Improving student writing in online and multimedia environments* (1st ed.). Jossey-Bass.

Englert, C. S. (1992). Writing instruction from a sociocultural perspective: The holistic, dialogic, and social enterprise of writing. *Journal of Learning Disabilities, 25,* 153–172.

Englert, C. S., Manalo, M., & Zhao, Y. (2010). I can do it better on the computer: The effects of technology-enabled scaffolding on young writer's composition. *Journal of Special Education Technology, 19*(1), 5–21.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

Graham, S. (2006). Writing. In P. A. Alexander & P. H. Winne, (Eds.) *Handbook of educational psychology* (2nd ed.; pp. 457–478). Erlbaum.

Graham, S., Harris, K. R., & Chambers, A. B. (2016). Evidence-based practices and writing instruction. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed.; pp. 211–266). The Guilford Press.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools: A report on the Carnegie Corporation of New York*. Alliance for Excellence in Education.

Grigal, M., Hart, D., Papay, C., & Smith, F., (2018). *Year two program data summary (2016–2017) of the TPSID model demonstration projects*. University of Massachusetts Boston, Institute for Community.

Grigal, M., Hart D., Smith, F.A., Domin, D., & Weir, C. (2017). *Think College National Coordinating Center: Annual report on the transition and postsecondary programs for students with intellectual disabilities (2014–2015)*. University of Massachusetts Boston, Institute for Community Inclusion.

Harris, K. R., Santangelo, T., & Graham, S. (2010). Metacognition and strategies instruction in writing. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 227–255). The Guilford Press.

Hayes, J. R., (2012). Modeling and remodeling writing. *Written Communication, 29,* 369–388.

Hayes, J. R., & Flower, L. S. (1987). On the structure of the writing process. *Topics in Language Disorders, 7,* 19–30.

Hayes, J. R., & Olinghouse, N. G. (2015). Can cognitive writing models inform the design of the common core state standards? *The Elementary School Journal, 115,* 480–497.

Hetzroni, O. E, & Shrieber, B. (2004). Word processing as an assistive technology tool for enhancing academic outcomes of students with writing disabilities in the general classroom. *Journal of Learning Disabilities, 37,* 143–154.

Higher Education Opportunity Act. (2008). Pub. L. No. 110-315. 122 STAT. 3078 (2008).

Hua, Y., Woods-Groves, S., & Yuan, C. (2018). Literacy interventions for young adults with intellectual and developmental disabilities in the inclusive postsecondary education settings: A review of a program of research. *Journal of Inclusive Postsecondary Education, 1, 1–20.*

Hughes, C. A., Morris, J. R., Therrien, W. J., & Benson, S. K. (2017). Explicit instruction: Historical and contemporary contexts. *Learning Disabilities Research & Practice, 32,* 140–148.

Hughes, C. A., Schumaker, J. B., McNaughton, D. B., Deshler, D. D., & Nolan, S. M. (2010). *The EDIT strategy.* The University of Kansas.

Jackson, L. G., Duffy, M. L., Brady, M. P., & McCormick, J. (2018). Effects of learning strategy training on the writing performance of college students with Asperger's Syndrome. *Journal of Autism and Developmental Disorders, 48,* 708–721.

Joseph, L. M., & Konrad, M. (2009). Teaching students with intellectual or developmental disabilities to write: A review of the literature. *Research in Developmental Disabilities, 30,* 1–19.

Kang, M., Ragan, B. G., & Park, J. H. (2008). Issues in outcomes research: An overview of randomization techniques for clinical trials. *Journal of Athletic Training, 43,* 215–221.

Karchmer-Klein, R. (2013). Best practices in using technology in support writing. In S. Graham, C. A., MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (2nd ed.; pp. 309–333). The Guilford Press.

Kernan, W. N., Viscoli, C. M., Makuch, R. W., Brass, L. M., & Horwitz, R. I. (1999). Stratified randomization for clinical trials. *Journal of Clinical Epidemiology, 52,* 19–26.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel.* Research Branch Report 8–75. Chief of Naval Technical Training: Naval Air Station Memphis.

Konrad, M., Trela, K., & Test, D. W. (2006). Using IEP goals and objectives to teach paragraph writing to high school students with physical and cognitive disabilities. *Education and Training in Developmental Disabilities, 41,* 111–124.

Lewis, R. B. (2000). Musings on technology and learning disabilities on the occasion of the new millennium. *Journal of Special Education Technology, 15,* 5–12.

MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research,* (2nd ed.; pp. 24–40). The Guilford Press.

McNaughton, D. B., & Hughes, C. A. (1999). *InSPECT: A strategy for finding and correcting spelling errors.* Edge Enterprises Inc.

Montgomery, D. J., Karlan, G. R., & Coutinho, M. (2001). The effectiveness of word processor spell checker programs to produce target words for misspellings generated by students with learning disabilities. *Journal of Special Education Technology, 16,* 27–42.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards.* Author.

Pennington, R. C. (2016). Write on! Using assistive technology and systematic instruction to teach sentence writing to students with moderate to severe disabilities. *Journal of Special Education Technology, 3,* 50–57.

Pennington, R. C., & Delano, M. E. (2012). Writing instruction for students with autism spectrum disorders: A review of literature. *Focus on Autism and Other Developmental Disabilities 27,* 158–167.

Philippakos, Z. A., MacArthur, C. A., & Coker Jr., D. L. (2015). *Developing strategic writers through genre instruction: Resources for grades 3–5.* The Guilford Press.

Regan, K., Evmenova1, A. S., Good, K., Legget, A., Ahn, S. Y., Gafurov, B., & Mastropieri, M. (2018). Persuasive writing with mobile-based graphic organizers in inclusive classrooms across the curriculum. *Journal of Special Education Technology, 33,* 3–14.

Schumaker, J. B., & Deshler, D. D. (2009). Adolescents with learning disabilities as writers: Are we selling them short? *Learning Disabilities Research & Practice, 24,* 81–92.

Schumaker, J. B., Nolan, S. M., & Deshler, D. D. (1985). *The error monitoring strategy: Instructor's manual.* The University of Kansas Center for Research on Learning.

Schunk, D. H., & Zimmerman, B. J., (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly, 23,* 7–25.

Silió, M. C., & Barbetta, P.M. (2010). The effects of word predication and text-to-speech technologies on the narrative writing skills of Hispanic students with specific learning disabilities. *Journal of Special Education Technology, 25,* 17–32.

Smith Lee, S. (2009). *Overview of the Federal Higher Education Opportunity Act Reauthorization.* Think

College Insight Brief, Issue No. 1. Institute for Community Inclusion, University of Massachusetts Boston.

Wong, B. Y. L. (2001). Commentary: Pointers for literacy instruction from educational technology and research on writing instruction. *The Elementary School Journal, 101,* 359–369.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement.* Riverside Publishing.

Woods-Groves, S., Hua, Y. Ford, J. W., & Neil, K. M. (2017). Efficacy of an electronic editing strategy with college students with intellectual and developmental disabilities. *Education and Training in Autism and Developmental Disabilities, 54,* 422–436.

Woods-Groves, S., Hua, Y., Therrien, W. J., Kaldenberg, E. R., Kihura, R. W., & Hendrickson, J. M. (2015). An investigation of the efficacy of an editing strategy with postsecondary individuals with developmental disabilities. *Education and Training in Autism and Developmental Disabilities, 50,* 95–108.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Received: 3 January 2019
Initial Acceptance: 26 February 2019
Final Acceptance: 18 May 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 158–172

App-Based Manipulatives and the System of Least Prompts to Support Acquisition, Maintenance, and Generalization of Adding Integers

Emily C. Bouck and Jiyoon Park
Michigan State University

Abstract: Virtual manipulatives are an emerging intervention to support students with disabilities in mathematics. Through a multiple probe across participants design, researchers examined use of an intervention package consisting of a virtual manipulative (i.e., the Two-Color Counter app-based manipulative) and the system of least prompts (SLP) to support students' acquisition, maintenance, and generalization of solving addition of integers problems. Researchers found a functional relation between the intervention package and middle school students' accuracy. All three students acquired the skill, with two able to consistently generalize as well as maintain adding integers. The three students were also generally independent when solving the problems with the app.

For students with disabilities, learning occurs across four stages: acquisition, fluency, maintenance, and generalization (Alberto & Troutman, 2009; Collins, 2012; Shurr et al., 2019). Students start the acquisition stage when they begin learning a new skill or concept (Collins, 2012; Snell & Brown, 2011). When students demonstrate the skill with 60% accuracy independently or 100% with support, they move into the fluency stage (Shurr et al., 2019). With fluency, students can be both accurate and efficient (Collins, 2012; Snell & Brown, 2011). During the fluency stage, the skill is not new and students become more independent. Once a student demonstrates a skill with sufficient accuracy (e.g., 60%) and realistic speed for the real world, the student moves into the maintenance stage (Shurr et al., 2019). Maintenance involves the performance of a skill when it is not proceeded by instruction (Alberto & Troutman, 2009). When students consistently demonstrate the skill without reteaching across time, they have achieved maintenance (Shurr et al., 2019). Generalization involves performing a skill consistently across different settings, people, or materials (Collins, 2012).

Research-supported instructional strategies that support acquisition include modeling; prompting systems, such as least-to-most prompting; task analysis; and visual supports (Snell & Brown, 2011; Shurr et al., 2019). Fading support and overlearning are two examples of instructional strategies that support maintenance of skills or concepts (Collins, 2012; Snell & Brown, 2011). Generalization is supported through distributed trials, natural reinforcement, and providing examples and non-examples (Collins, 2012; Snell & Brown, 2011; Shurr et al., 2019).

Mathematics Education for Students with Disabilities

As with other areas, mathematics learning for students with disabilities follows the stages of acquisition, fluency, maintenance, and generalization. However, researchers suggest limited attention is paid to issues of maintenance and generalization with regards to academic content for students with disabilities (Spooner et al., 2012). Although researchers found mathematics interventions and instruction positively benefit student acquisition of mathematical content (cf., Barnett & Cleary, 2015; Hudson et al., 2018; King et al., 2016; Spooner

Correspondence concerning this article should be addressed to Emily C. Bouck, 343a Erickson Hall, 620 Farm Lane, Michigan State University, East Lansing, MI 48824. E-mail: ecb@msu.edu

et al., 2018), the mathematical learning of students with disabilities needs to extend beyond acquisition. The ability to maintain and generalize allows students with disabilities (a) access to mathematical content that builds upon previous content, (b) independence in mathematics and content areas that involve math, and (c) to perform mathematics-based independent living skills (e.g., purchasing, cooking; Alwell & Cobb, 2009; Geary et al., 2017; Powell et al., 2013).

Manipulatives. Researchers exploring mathematics interventions for students with disabilities identified interventions and instructional strategies that support acquisition. One such intervention is manipulatives (Bouck & Park, 2018; Spooner et al., 2018). Manipulatives are concrete or virtual objects students can manipulate to aid in understanding mathematical concepts (Bouck, Working, & Bone 2018). Concrete manipulatives are physical objects while virtual manipulatives are digital objects available via the Internet or as an app (Bouck, Working & Bone, 2018). Researchers found the use of manipulatives to be an evidence-based or research-supported practice for students with disabilities (Bouck & Park, 2018; Spooner et al., 2018). However, the majority of research involves concrete manipulatives; use of virtual manipulatives as a tool to support students with disabilities is still emerging (Bouck & Park, 2018).

Within the limited research base regarding virtual manipulatives to support mathematics for students with disabilities, the results are positive. In a single case design study, Satsangi and Bouck (2015) found three secondary students with learning disabilities acquired—as well as maintained and generalized—accuracy related to area and perimeter when using a virtual manipulative. Bouck et al. (2019) found three middle school students with disabilities learned to solve division with remainders problems through the use of an app-based manipulative: Cuisenaire® Rods. Much of the other research regarding virtual manipulatives and students with disabilities involves the use of the tool within a graduated system of instruction (i.e., the virtual-representational-abstract [VRA] instructional sequence or the virtual-abstract [VA] instructional sequence; Bouck, Bassette, et al., 2017; Bouck, Park, et al., 2018; Bouck, Park, et al., 2017) or comparing virtual and concrete manipulatives (Bouck et al., 2014; Bouck, Chamberlain, & Park, 2017; Bouck, Shurr, et al., 2018). With regards to virtual manipulatives as part of an instructional sequence, the existing research supports the acquisition of various mathematical content by middle school students with disabilities (e.g., basic operations, fractions, algebra), but results are more mixed with regards to students maintaining after intervention ends (Bouck, Bassette, et al., 2017; Bouck, Park, et al., 2018; Bouck, Park, et al., 2017). Researchers comparing virtual and concrete manipulatives for students with disabilities found both types of manipulatives successful across mathematical content (e.g., subtraction, fractions), with some differences reported with regards to independence and preference, in favor of virtual manipulatives (Bouck et al., 2014; Bouck, Chamberlain, & Park, 2017; Bouck, Shurr, et al., 2018).

Explicit instruction. Researchers also noted explicit instruction as an evidence-based practice for teaching mathematics to students with disabilities (Doabler & Fien, 2013; Riccomini et al., 2017; Spooner et al., 2018). With explicit instruction, an instructor first models a few problems, then the student attempts a few problems while the instructor provides prompts or feedback as needed, and finally the student completes problems independently (Doabler & Fien, 2013). Much of the research regarding virtual manipulatives—as well as concrete manipulatives—involves explicit instruction as the mode of delivery (Bouck & Park, 2018; Bouck, Satsangi, & Park, 2018). Both the virtual and the concrete graduated instructional sequences (i.e., the VRA and VA as well as the concrete-representational-abstract instructional sequences) employ explicit instruction (Bouck, Bassette, et al., 2017; Bouck, Park, et al., 2018; Bouck, Park, et al., 2017; Bouck, Satsangi, & Park, 2018). Researchers have also successfully taught students with disabilities to use virtual manipulatives outside of an instructional sequence via explicit instruction (Bouck et al., 2019; Satsangi et al., 2018).

Prompting. Prompting—such as the system of least prompts (SLP)—is a form of systematic instruction, which is supported as an evidence-based or research-based practice with regards to mathematics instruction and students with moderate or severe disabilities

(Browder et al., 2008; Spooner et al., 2018). The SLP involves a hierarchy of prompts that are applied in increasing intensity or intrusiveness if a student fails to respond (Shepley et al., 2018). Generally, Shepley et al. (2018) found use of the SLP as a stand-alone intervention to be an evidence-based practice for students with disabilities as well as support for this approach to teach students to use technology. They also suggested its use in academics was appropriate for secondary students.

Current Study

Given research suggesting such effective interventions for students with disabilities as explicit instruction, the SLP, and manipulatives (Bouck & Park, 2018; Shepley et al., 2018; Spooner et al., 2018), the researchers sought to explore the impact of an intervention package to support students across the learning stages. The researchers felt attention to delivery of explicit instruction regarding use of app-based manipulatives to promote understanding of addition of integers, supported by the SLP, was a need in the field given much of the research on virtual manipulatives involves the tool as part of an instructional sequence. The researchers hypothesized an intervention package as opposed to just one intervention (e.g., virtual manipulatives or explicit instruction or SLP) would result in improved acquisition and maintenance for the population in question.

In this study, researchers sought to explore the functional relation between the intervention package of use of an app-based virtual manipulative and the system of least prompts and middle school students' accuracy and independence in solving mathematical problems. The study sought to answer the following research questions: (a) Does a functional relation exist between the intervention package (i.e., use of a virtual manipulative and the SLP) and student accuracy? (b) with what percentage of independence will students complete addition of integers problems? (c) Are students able to generalize solving problems when no virtual manipulative or prompts are available? and (d) what are student perceptions of the use of the virtual manipulatives in solving mathematical problems?

Method

Participants

The researchers included three middle school participants in this study. All three of the seventh and/or eighth-grade students received their mathematics instruction in a special education class taught by a special education teacher. Two of the students received their mathematics instruction during one class period and the other two during a different period. Researchers worked individually with students during their 61-minute period mathematics class or the shorter 30-minute intervention period, which the special education teacher used for life skills. The special education teacher selected students for participation in the study based on her assessment of their mathematical struggles and perception of the benefits of one-on-one instruction. The students' mathematical struggles were evaluated via the KeyMath-3 assessment.

Nick. Nick was an eighth-grade student, who was 14-years-old at the time of the study. Nick, a white male, was determined to be eligible to receive special education services under the category of learning disability. Per his Individualized Education Program (IEP), Nick met the state criteria for learning disability (i.e., as determined by the ability-achievement discrepancy) in the area of mathematics calculation, reading, and reading comprehension. Nick also had a secondary disability of ADHD. Nick's IQ, per his most recent evaluation, was 71. On the most recent mathematics achievement data available within Nick's folder, he was in the fifth percentile on the Test of Early Mathematics Ability-3 (TEMA-3; Ginsburg & Barroody, 2003). On the researcher-administrated KeyMath-3, Nick's numeration score was 23, which was a 4.8 grade equivalency and his total operations scores was 39, a 3.6 grade equivalency.

Jess. Jess was a seventh-grade, 14-year-old, white female. Jess received special education services under the category of autism. Jess was primarily educated in the self-contained special education classroom, taught by a special education teacher, which included her mathematics class. Jess received special education services under the category of autism, although no autism assessment scores were pro-

vided within her file. Her Individualized Education Program (IEP) just indicated the category of autism. On the Keymath-3 administered by researchers, Jess's numeration score was 15, a 2.5 grade equivalency. Her total operations score was 25, a grade equivalency of 2.6. As reported by the teacher, Jess demonstrated more advanced mathematics skills than she tested.

Cece. Cece was a seventh-grade, 13-year-old, white female. Cece was eligible for special education services under the category of Other Health Impairment (OHI). According to Cece's records, she was eligible for OHI due to a condition called Chromosome3Q29 microduplication syndrome. Cece was evaluated for autism, but determined not to meet the criteria due to no delays in communication and a lack of preoccupation with parts of an object. The Autism Diagnostic Observation Schedule (ADOS-2; Lord et al., 2012) assessment within her file suggested no signs of autism (i.e., score of 3). Cece's file indicated an IQ of 70. On the researcher-administrated Keymath-3, Cece's numeration score was 16 and her total operations score was 29, a 2.8 and 2.9 grade equivalency, respectively.

Setting

Researchers conducted the study in a public middle school in a small Midwestern town located about half-an-hour from a large research university and the state capital. The school district consisted of four schools: two elementary schools, one middle school, and one high school. The student enrollment across the district was about 2000 students, with just over one-fourth occupying the middle school. The middle school educated students in grades sixth through eighth. The student population was predominantly white, with less than ten percent African American, Latinx, and/or two or more races. Under one-third of the student population in the district qualified for free and reduced lunch and approximately 10% had an IEP.

In terms of the setting for the study itself, the researchers worked with students at a table in a hallway outside of their special education classroom. Researchers worked with students during either their mathematics class or the schoolwide intervention period. Each session was generally less than 10 minutes in length; pre-training was longer. The hallway was generally quiet during the time of the study as all other classes were in session.

Materials

Researchers used three main materials in the study: the iPad, the Brainingcamp Two Color Counter app, and probes. The app—a virtual manipulative—was used on the iPad. The app consisted of two different colored counters—yellow for positive marked with a plus sign and red for negative marked with a negative sign. The researchers used probes in all phases of the study and provided the dependent variable. Each probe reflected the mathematical focus of this study, adding integers (e.g., $3+ -5$ or $-2 +8$). Each probe consisted of five problems. The researchers presented the problems the problem horizontally in a 2×3 table-like format with one problem per "cell." To develop the probes, researchers created all possible integer problems involving numbers from -10 to 10, excluding the number 0. The researchers then randomly assigned each problem to assessments; they repeated no problems across assessments.

Independent and Dependent Variables

The independent variable for the study was the intervention package of a virtual manipulative (i.e., app-based manipulative) and the system of least prompts. During each intervention session, researchers allowed students to use the app-based manipulative of Two-Color Counters (Brainingcamp, 2018) to represent and model positive and negative integers and researchers employed the SLP, which consisted of independently solving the problem, gesture, indirect verbal, direct verbal, and modeling. The researchers applied the SLP to the task analysis steps to solve the addition of integers problems and activated it when a student did not initiate the next step within 10 seconds and/or completed the step incorrectly, up until the last step of answering the problem correctly. Although students could be prompted to use the virtual manipulatives or check the sign of the counters left, they were allowed to write down an incorrect answer.

Two dependent variables were included in the study: accuracy and independence. Accuracy was defined as the percentage of problems—out of five—that students solved correctly on each probe delivered during each phase of the study. Independence was determine based on the percentage of task analysis steps students completed independently—without any prompts—during each session. For each problem of each assessment, nine steps existed to solve, resulting in a total possible of 45 steps per session (task analysis available upon request from first author).

Experimental Design

Researchers employed a single case multiple probe across participants design. Each participant began baseline simultaneously and completed five baseline sessions prior to the first participant beginning intervention. After the first student obtained a stable baseline—meaning 80% of his/her baseline data fell within 25% of the median of that data—and had a zero-celeration or deceleration trend, the student entered the pretraining phase. During pretraining, the researchers taught students to how to solve addition of integer problems using the app-based manipulative, using the task analysis. Once a student achieved at least 90% accuracy *and* independence on the pretraining probe, the first student began the intervention phase. If the student did not achieve 90% for accuracy *and* independence, researchers provided another pre-training session.

During each session of the intervention, the student completed a generalization probe first (i.e., without the app) and then an intervention probe with the use of the app-based manipulative. Once Student 1 completed three intervention sessions in which s/he achieved 80 or 100% accuracy *and* independence on the three intervention probes, Student 2 began pre-training and then moved into intervention. Similarly, Student 3 entered pretraining and intervention under the same process—Student 2 obtained 80–100% accuracy *and* independence on three intervention probes. While Student 1 was in intervention, Students 2 and 3 still completed baseline sessions and when Student 2 entered intervention, Student 3 completed additional baseline probes.

Procedure

Data collection occurred one-on-one with students, with the exception of when two researchers were present to collect interobserver agreement and fidelity of implementation data. Session lengths ranged from less than one minute (baseline) to 5:40 minutes (intervention). Each student participated in a minimum of five baseline sessions, five intervention sessions, five generalization probes during the intervention phase, and four longer-term maintenance sessions. Two researchers completed data collection. The first—the first author—was a special education professor with years of experience studying mathematics interventions for students with disabilities. The other researcher was a doctoral candidate who had three-plus years working with the faculty member researching mathematical interventions for students with disabilities.

Baseline. Each student completed a minimum of five baseline sessions. During each baseline session, students completed a probe. Each probe consisted of five addition of integer problems; for each problem there was one positive and one negative number. Researchers printed all probes on an 8.5×11 piece of white paper and students used a pencil to solve the problems. Students received no instruction or prompts during baseline, although the iPad with the app-based manipulative was set in front of students and students were told they could use the virtual manipulative to aid in solving. Researchers recorded only accuracy data during baseline.

The first student left baseline after s/he completed five baseline sessions and his/her baseline data were stable with a zero-celeration or deceleration trend. The second student left baseline once the first student achieved 80% accuracy *and* independence on three intervention probes and his/her baseline data were stable with a zero-celeration and deceleration trend. Similarly, the third student left baseline when the second student's intervention probe accuracy data were 80% or greater for accuracy *and* independence on three intervention probes and the

third student's baseline data were also stable with a zero-celeration or deceleration trend.

Pre-training. After each student met the criteria to transition out of baseline, s/he entered pretraining. During the pre-training phase, researchers individually trained each student on how to use the app-based virtual manipulative—the Two-Color Counter (Brainingcamp, 2018). To train students, the researcher used explicit instruction, meaning the researcher modeled two addition of integer problems using the app, modeling and verbally narrating as a means to teach how to use the app. Next, the researcher guided the student as s/he worked to solve two problems using the app. Last, the researcher allowed the student to solve five problems independently with the app-based manipulative. If the student achieved over 90% on the independent portion for both accuracy and independence, when using the app-based manipulative, the student moved onto intervention. If not, the student repeated the pre-training phase at the following session. Note, all students achieved the pre-training criterion in one session. The researchers set a high mastery criterion for the pre-training as it was the only session in which explicit instruction was provided.

To train students with the Two-Color Counter manipulative, researchers explained the yellow chips represented positive numbers and highlighted the positive—or plus—symbol on the yellow chips. Similarly, the researcher explained the red chips represented negative numbers and noted the negative symbol on the red chips. During the modeling portion of the explicit instruction pre-training, the researcher physically demonstrated how to use the app and move chips onto the work space, while verbally narrating (i.e., providing a think aloud). With the problem $4+ \ -6 = \_$, for example, the researcher read the problem and then stated how the four is a positive integer and thus could be represented by four of the yellow chips, and subsequently pulled four yellow chips onto the work space of the app. The researcher then stated that the 6 was a negative six, meaning 6 less than zero and could be represented by 6 of the red chips, and subsequently pulled out six red chips. The researcher explained that a positive one, meaning one more than zero and a negative one, meaning one less than zero, when added together equal zero and showed that when a negative red chip is moved on top of a positive yellow chip, they both grayed in color, indicating a sum of zero. The researcher then demonstrated this for the remaining three positive yellows chips, until only two negative red chips were left. The researcher stated, "looking at our work space we can now conclude that $4+ \ -6 = -2$" and wrote that answer on the piece of paper. The same procedures were used regardless of the problem – whether, as noted $4+ \ -6$ or $7+ \ -3$ or $-9 \ +5$ or $-2 \ +8$.

Intervention. Each student completed a minimum of five intervention sessions. During each intervention session, researchers delivered an intervention package consisting of the app-based manipulative (Two-Color Counters; Brainingcamp, 2018) and the SLP. Researchers opened the app on the iPad and had it ready for the start of the intervention session. Researchers instructed students to use the app to aid in solving the problems. In terms of the SLP, researchers used a prompting hierarchy consisting of independent, gesture (e.g., pointing to or gesturing towards the app or a feature on the app), indirect verbal (e.g., stating, "what do you next"), direct verbal (e.g., stating, "you need to make sure you have the number of negative red chips on the work space that is written in the problem), and modeling (e.g., physically moving a red negative chip onto a yellow positive chip to cancel it out and reflect a sum of zero). Researchers worked through the SLP for each step of the task analysis for solving addition of integer problems and implemented a higher level of the prompting hierarchy when the student did not initiate action within 10 seconds (i.e., researchers used a 10-second time delay). Researchers allowed students to write down an incorrect answer, as long as they followed the steps of the task analysis.

The task analysis steps for solving addition of integer problems consisted of nine steps, resulting in a total of 45 steps per session. These steps included: (a) determining if the first number of the problem was positive or negative, (b) dragging the number of red or yellow chips for the first number in the problem onto the white space, (c) determining if the second number of the problem was positive or negative, (d) dragging the number of

red or yellow chips for the second number in the problem onto the white space, (e) making pairs of one yellow/positive and one red/negative, (f) counting the number of unpaired objects, (g) determining the sign of the counters remaining, (h) writing down the answer, and (g) clearing the screen.

Maintenance. Students completed two maintenance sessions at both two and four weeks after their last intervention session. During the maintenance probes, researchers gave students access to the Two-Color Counters virtual manipulative (Brainingcamp, 2018) to determine if students maintained solving addition of integer problems with use of the tool.

Generalization. Researchers administered generalization probes during the intervention phase. Each generalization probe was similar to baseline and intervention probes (i.e., five problems), but researchers did not allow students access to the app-based manipulative when solving the problems and did not implement the SLP. Generalization probes occurred the same day as intervention and maintenance probes; students completed the generalization probes before completing the intervention or maintenance probes.

Interobserver Agreement and Procedural Fidelity

Interobserver agreement data (IOA) accuracy data were collected by researchers for at least 20% of the baseline, intervention, maintenance, and generalization sessions. During the intervention sessions, researchers also gathered IOA data relative to independence. To determine IOA—for accuracy or independence, researchers summed the number of agreements of the respective dependent variable and divided that by the sum of the number of opportunities for agreement or disagreements, and finally multiplied the quotient by 100 for a percentage. IOA was 100% for accuracy and independence for each participant.

The second researcher assessed for fidelity of implementation data during 40% of intervention session by completing a checklist as to the first researchers' procedural fidelity. The second researcher checked that the student (a) was given the iPad app; (b) was prompted only after a 10-second delay for each step of the task analysis; and (c) was given prompts in the set prompting hierarchy order, as needed. Procedural fidelity was 100% for each participant.

Social Validity

To assess for social validity, researchers interviewed both the participating students after the intervention ended. Each of the students answered what they liked and did not like about learning mathematics with the app-based manipulative as well as what they thought about the researcher helping them if they got stuck. Students were also asked if they would continue to use an app-based manipulative if allowed in their classroom and if they felt the manipulative helped them to learn the mathematics.

Data Analysis

To analyze the accuracy data, researchers conducted both a visual analysis well as computed various metrics for the accuracy dependent variable. Specifically, researchers analyzed for immediacy of effect, level, trend, and effect size (see Table 1). Researchers visually examined the graphed data to determine the immediacy of the intervention effect, comparing the last baseline for each student to his/her first intervention session. For level, researchers calculated the median for baseline and intervention—separately—and then determined if 80% of the data for each phase, respectively, fell within a 25% range of the median (Gast & Spriggs, 2014). For trend, researchers determined the middle point, mid-rate, and mid-date for both the baseline and intervention data, separately (White & Haring, 1980). Using the graphs, the researchers then marked these points and drew a line connecting the mid-rate and the mid-date. If the line for that phase was straight, the trend was zero-celerating, if it was sloping up it was accelerating, and if the line was sloping down the trend was decelerating.

For effect size, the researchers calculated PND and Tau-U for accuracy. To calculate PND, the researchers determined the highest baseline data point. The number of intervention data points higher than the highest baseline data point, divided by the number of intervention data points (e.g., 5), and then multiplied by 100 was the PND. To calculate

TABLE 1

Accuracy Data Analysis Summary of the App-Based Manipulative and Prompting

Measure	*Baseline*	*Intervention: Virtual Manipulatives + Prompting*	*Generalization*	*Maintenance*
		Nick		
Range	0	5 (100%)	3–5 (60–100%)	5 (100%)
Mean	0	5 (100%)	4.6 (92%)	5 (100%)
Median	0	5 (100%)	5 (100%)	5 (100%)
Stability	stable	stable	stable	stable
Trend	zero-celerating	zero-celerating	zero-celerating	zero-celerating
Tau-U[a]	—	1.0	1.0	1.0
PND	—	100%	100%	100%
# of sessions	5	5	5	4
		Jess		
Range	0	5 (100%)	0–5 (0–100%)	4–5 (80%–100%)
Mean	0	5 (100%)	3.6 (76%)	4.2 (85%)
Median	0	5 (100%)	5 (100%)	4 (80%)
Stability	stable	stable	stable	variable
Trend	zero-celerating	zero-celerating	decelerating	accelerating
Tau-U[a]	—	100%	0.8	1.0
PND	—	100%	80%	100%
# of sessions	8	5	5	4
		Cece		
Range	0	3–5 (60–100%)	0–3 (0–60%)	1–3 (20–60%)
Mean	0	4.4 (88%)	1.6 (30%)	2 (40%)
Median	0	5 (100%)	2 (40%)	2
Stability	stable	variable	variable	variable
Trend	zero-celerating	accelerating	accelerating	zero-celerating
Tau-U[a]	—	1.0	0.8	1.0
PND	—	100%	80%	100%
# of sessions	9	5	5	3

Note: [a] Tau-U conducted between baseline and intervention, between baseline and generalization, and then baseline and maintenance.

Tau-U, researchers used an online calculator, inputting both the baseline and intervention data points, for each of the dependent variables separately (Vannest et al., 2016). The online Tau-U calculator then produced a Tau-U percentage. Applying accepted metrics, a Tau-U greater than .80 represented a very large effect, one between .60 and .80 a large effect, and one between .20 and .60 a moderate effect (Vannest & Ninci, 2015).

Researchers analyzed the mean and range of the independence data across the five intervention sessions. Researchers also analyzed the level of prompts provided to students—determining if those prompts were gesture, indirect verbal, direct verbal, or modeling.

Results

For each of the three participants, a functional relation was established between the dependent variable of accuracy and the intervention package of an app-based manipulative and prompting (see Figure 1). The students were also generally independent during each intervention session, needing few prompts to complete the session with the app-based manipulative.

Nick

During baseline, Nick answered zero questions correctly for each of the five sessions, resulting in a stable and zero-celerating base-

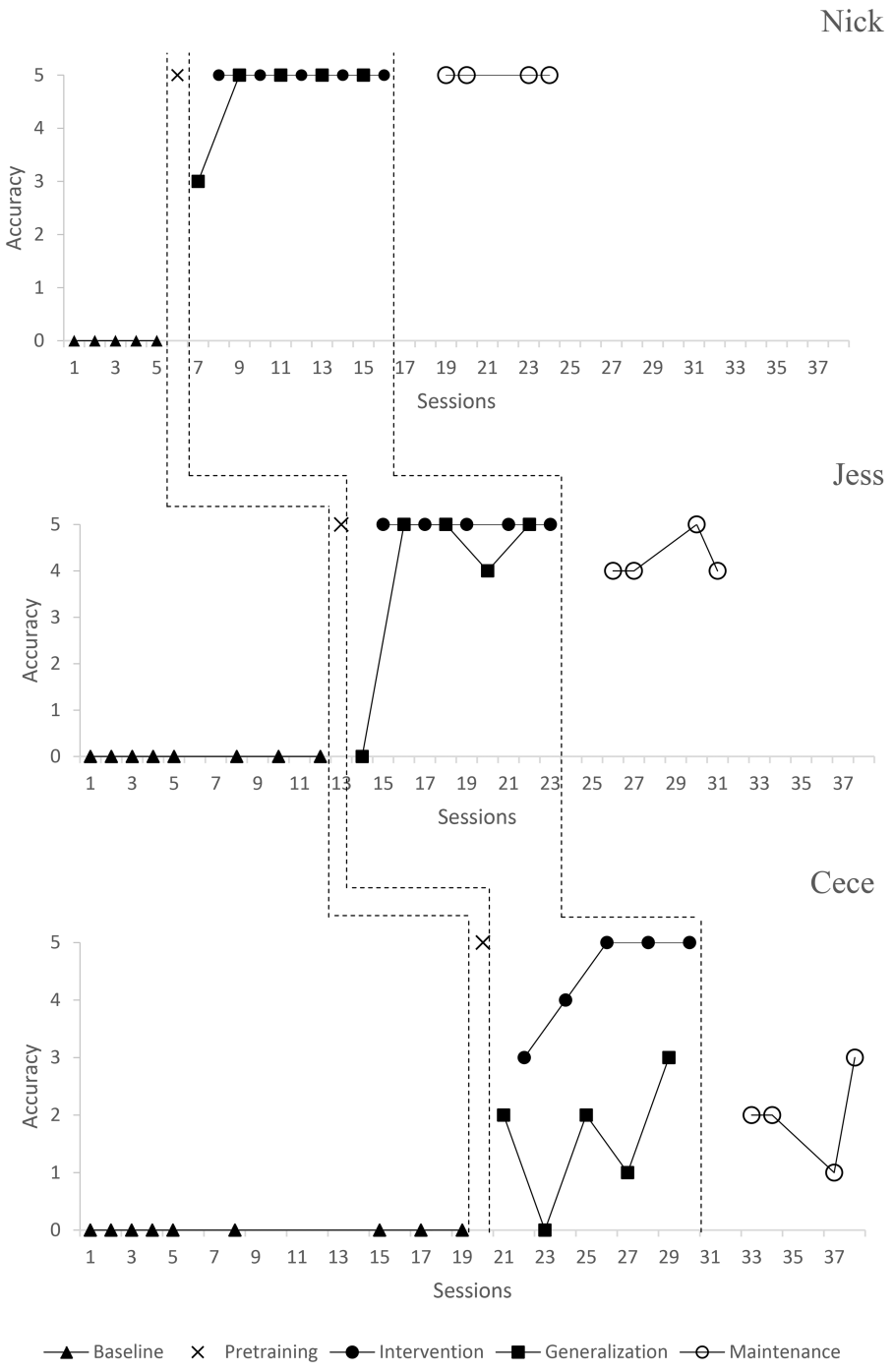


Figure 1. Accuracy of Solving Addition of Integers Problems During Acquisition, Generalization, and Maintenance.

line phase. When trained on using the Two-Color Counters (Brainingcamp, 2018) app-based manipulative, Nick scored 100% on the independent phase of explicit instruction. Nick maintained 100% during each of his five intervention sessions, resulting in a stable and zero-celerating trend intervention phase. The Tau-U for his intervention phase was 1.0, a very strong effect, and his PND with baseline was 100%. Nick scored 80% on the first generalization session and 100% for each of the other four sessions, which resulted in a stable and zero-celerating phase. As compared to the baseline phase, the Tau-U for generalization was 1.0 and the PND was 100%. Nick also scored 100% on each of his four maintenance sessions, also resulting in a Tau-U of 1.0 and a PND of 100% for this phase.

Nick was generally independent during each of his five intervention sessions. On the first intervention session, Nick's rate of independence was 95.6%, meaning he needed two prompts, both of which were gestures and delivered on the first task analysis step—if the first number was positive or negative. For each of the next four sessions, Nick was 100% independent.

Jess

Jess's accuracy for each of her eight baseline sessions were 0. She too, scored 100% on each of her intervention sessions, resulting in a Tau-U of 1.0—a very strong effect—and a PND of 100%. Her intervention data were stable and zero-celerating, as were her baseline data. As with Nick, Jess was trained on the app-based manipulative in one session, scoring 100% on the independent portion and demonstrating mastery. Jess's first generalization session score was 0, after which she scored 80 or 100%. Her generalization data were stable with a decelerating trend, although the Tau-U was 0.80 and the PND was 80%. Across the four maintenance sessions, Jess scored 80 or 100%. The Tau-U was 1.0 and the PND was 100% for her maintenance phase.

Jess was also independent during intervention. Like Nick, Jess needed prompts during her first intervention session, but was 100% independent for the subsequent four intervention sessions. Specifically, during her first intervention session, Jess needed four prompts, all gestures. Two were for determining if the first number was positive or negative, one for making pairs with the positive and number, and the last for finding the sign of the answer.

Cece

Cece answered zero questions correctly during baseline for each of her nine sessions. Although she achieved mastery criteria with the app-based manipulative training in one session (100%), her accuracy during intervention was more gradual. She answered three correct on the first intervention, four on the second, and then five for the last three sessions, resulting in a variable but accelerating trend. Her PND for intervention was 100% and her Tau-U was 1.0. Her generalization data were also variable, but she never answered more than three problems correctly during generalization. The Tau-U for her generalization data was 0.8 and her PND was 80%, with a range of zero-to-three correct answers. During maintenance, her accuracy was variable, ranging from one to three correct answers.

Cece's independence was more variable, which ranged from 80.6%–96.3%. Cece never completed an intervention session without at least one prompt. While she gradually improved her independence from session 1 to session 4 (range 90–96.3%), it was her last session in which she needed the most prompting (independence was 80.6%). Researchers gave Cece prompts at the gesture, indirect verbal, and direct verbal levels, and those were spread across the different steps of the task analysis, including determining the sign of the first number, determining the sign of the answer, making pair, clearing the app screen, and writing down the answer.

Social Validity

The students all expressed a preference for using the app-based manipulative on the iPad. It was most pronounced with Cece, who struggled during the generalization phases when she was told she was not going to use the iPad. She would repeatedly ask before starting the generalization probe as well as state it would be easier and better with the iPad while she

was completing the problems. Nick and Jess also preferred to use the app-based manipulative, as they enjoyed using the iPad. However, they did report they were faster when doing the math without the iPad. They all stated they would use an app to solve other types of math problems.

Discussion

This study focused on the use of an intervention package consisting of a virtual manipulative (i.e., the Two-Color Counter app-based manipulative) and the SLP to support students' acquisition, maintenance, and generalization of solving addition of integers problems. In this study, researchers found a functional relation between the intervention package and middle school students' accuracy. All three students acquired the skills, with two able to consistently generalize as well as maintain adding integers. The three students were also generally independent when solving the problems with the app; two students achieved 100% independence while one continued to need prompts throughout the five intervention sessions.

All three students acquired the skill of addition of integers when using a virtual manipulative and supported by prompting, as suggested by all three achieving 100% accuracy during intervention. For Nick and Jess, this was immediate and consistent, while it was more gradual for Cece. As such, this study supports the limited but emerging research regarding the efficacy of virtual manipulatives. Previous researchers found secondary students with disabilities successful in acquiring mathematical content when using virtual manipulatives (Bouck et al., 2019; Satsangi & Bouck, 2015; Satsangi et al., 2018). To date, researchers demonstrated the success of virtual manipulatives as an intervention for secondary students with disabilities across the mathematical domains of area and perimeter, linear algebra, division with remainders, and now adding integers (Bouck et al., 2019; Satsangi & Bouck, 2015; Satsangi et al., 2018).

Although existing literature presents mixed results regarding students' maintaining their learning with virtual manipulatives and virtual manipulative-based instructional sequences (cf., Bouck, Park, et al., 2018; Satsangi & Bouck, 2014), in this study, students maintained their learning, at least at rates higher than baseline. When given access to the virtual manipulative, but no prompting, Nick and Jess maintained at rates of 80% both two- and four-weeks post- intervention. As with acquisition, Cece's maintenance data were more variable, but still above baseline levels. Given that a particular instructional strategy targeting maintenance was not incorporated into the study (see Shurr et al., 2019), researchers interpreted the maintenance data positively.

Students in this study also generalized their learning, meaning they solved addition of integer problems without a virtual manipulative, prompting, or any explicit instruction on how to solve the problems without the manipulative. These results were more true for Nick and Jess, who both achieved 100% accuracy on generalization sessions. For Cece, however, generalization was less successful; her highest accuracy during a generalization session was 60%. Yet, Cece's generalization data are not surprising as she struggled more with intervention when using the manipulative, she needed more prompts to successfully complete the task analysis steps during intervention, and she repeatedly said during the generalization sessions that she wanted to use the app and asked why she could not. As with maintenance, specific instructional strategies targeting generalization were not implemented within the study (see Shurr et al., 2019), but students applied what they learned about solving the addition of integers problems with the app-based manipulative to solving the problems abstractly.

In terms of independence, Jess and Nick were generally independent – each achieving 100% independence and hence needing no prompts during intervention sessions. In contrast, Cece needed prompts during each intervention session. Although Cece displayed gradual improvement in terms of independence during her first four interventions sessions, she needed the most prompting during session five. Cece's need for more prompting may have likely contributed to her lower maintenance and generalization scores.

Integer Instruction

Within the literature regarding mathematics interventions for students with disabilities, limited literature examines instruction in integers; yet, for secondary students an understanding of integers can support algebra as well as aspects of daily living (e.g., temperature or money; Gagnon & Maccini, 2001). Maccini and Hughes (2000) and Maccini and Ruhl (2000) both explored problem representation and problem solution of middle school students with learning disabilities with regards to solve integer problems. Through use of the concrete-representational-abstract instructional sequence along with a word problem strategy, Maccini and Hughes (2000) and Maccini and Ruhl (2000) both found students were able to solve the problems as well as maintain their success with solving the problems. Similar to those studies, this study emphasized the use of a manipulative (i.e., virtual vs. concrete) and explicit instruction. However, this study focused on computation; future research should focus on teaching addition—or subtraction—of integers via word problems and in natural contexts (e.g., temperature, time, money).

In this study, researchers used two-color counters (or chips) to symbolize positive and negative numbers, which may reflect a neutralization model (Battista, 1983; Stephan & Akyuz, 2012). While the use of the two-color counters was effective in this study, researchers focused solely on addition, and with smaller numbers. Another model approach for working with integers involves the use of a number line (Stephan & Akyuz, 2012). Researchers suggested a number line model presents advantages (Stephan & Akyuz, 2012); it may also lend itself easier to subtraction of integers as well as help students to see the real-world application of integers. Future researchers may seek to utilize a virtual number line as a tool for teaching students with disabilities addition and subtraction of integers.

Implications for Practice

In terms of implications for practice, the results of this study support the intervention package of virtual manipulatives and the system of least prompts as an effective intervention, both in terms of accuracy as well as independence. Further, there was support for the efficiency of the intervention. After one training session involving explicit instruction, students successfully completed five intervention sessions. From this, the students maintained, at higher levels than baseline, two and four weeks after the intervention ended. Similarly, the students generalized, most often at rates higher than baseline, when they did not have the virtual manipulative. Further, each session was relatively short in duration. While the training session was obviously longer as the researcher modeled two problems and then supported two problems, intervention session lengths ranged from 0:31 to 4:11 for Nick, 1:35 to 4:29 for Jess, and 4:02 to 8:13 for Cece. Given the efficiency, in addition to the efficacy, teachers may want to consider how such an intervention package can be used in their classroom with their students.

An implication of use however, is that teachers need the technology to implement the intervention package. While the app-based manipulative itself was relatively inexpensive (e.g., $0.99), teachers must also purchase the device it goes on. While researchers used an iPad, app-based manipulatives are available from Chromebooks or computers. When virtual manipulatives are not an option, teachers may consider concrete manipulatives. To date, researchers have yet to compare concrete and virtual two-color counters. However, a difference between the virtual manipulative and the concrete is that the virtual manipulative counters come with positive and negative symbols, whereas the concrete ones do not. Researchers may seek to compare concrete to virtual two-colored counters, with a particular lens to differences that may contribute to one being more effective. Teachers could also make use of a virtual, concrete, or representational number line; previous researchers suggested the number line model is more advantageous (Stephan & Akyuz, 2012). Researchers may seek to compare concrete, virtual, and representational (i.e., paper-based drawing) number lines for use to support addition—and subtraction—of integers.

A final implication involves that students with developmental disabilities were able to acquire—as well as generalize and maintain—

more grade-level appropriate content. Exposure to the concept of integers operations generally occurs in middle school. Through this study, with its approach of an app-based manipulative and the SLP, students were exposed to grade-level or closer to grade-level content. It is important to consider appropriate grade-level content for students with developmental disabilities (Spooner et al., 2018).

Limitations and Future Directions

This study, as all research, presents limitation. One limitation of the study is that researchers were unable to separate the impact of the virtual manipulative from the support of the SLP. While the students generally needed few prompts—with two achieving 100% accuracy—it is still difficult to untangle the effectiveness of each. While one might hypothesize that Nick and Jess would have successfully acquired adding integers without the support of prompting, the case is not as clear for Cece. Another limitation was that Cece struggled to a greater extent than Nick and Jess, including her last intervention session at 80.6% independence. Researchers may have wanted to extend the intervention to determine if the low independence, while still high accuracy, was a one-time situation or a pattern, as previously Cece's rate of independence was increasing across sessions one through four. Cece also struggled to a greater extent with generalization and maintenance, which researchers hypothesized was related to her lower rates of independence as well as her repeated verbalized desire to only use the app-based manipulative and dislike the generalization sessions.

Additionally, a limitation involves that researchers did not probe for generalization after a period of time. While researchers were able to determine students generalized the skill—at least a small amount in the case of Cece—during intervention, they failed to probe for generalization during maintenance. Ideally, the students would have not only maintained their accuracy solving addition of integers problems when provided access to the virtual manipulative, but also without. Researchers should seek to collect generalization data at later points in future research, especially when high rates of generalization during intervention are found.

Two final limitations involve participants and setting or mode. One might determine that a limitation was the difference in disability classifications of the students in this study. However, all students were served in the same self-contained special education classroom by the same teacher and all struggled with adding integers. Also, Nick's IQ was 71, which according to the American Association on Intellectual and Developmental Disabilities (n.d.,), an IQ under 75 can still quality as an intellectual disability, with concurrent adaptive behavior limitations. It may be that Nick received the diagnosis of learning disability as a result of the school's reluctance to label as a mild intellectual disability (Gargiulo, 2012; Polloway, Lubin, Smith, & Patton, 2010). Likewise, one might determine a limitation was that intervention sessions were all delivered one-on-one with a member of the research team in the hallway, as opposed to in the classroom by the teacher and/or in a group setting. Researchers should seek to implement similar intervention within a classroom setting, with the teacher as the interventionist, and/or delivering the intervention to more than one student at a time.

References

Alberto, P., & Troutman, A. (2009). *Applied behavior analysis for teachers* (8th ed.). Pearson.

Alwell, M., & Cobb, B. (2009). Functional life skills curricular interventions for youth with disabilities: A systematic review. *Career Development and Transition for Exceptional Individuals, 32,* 82–93. doi:10.1177/0885728809336656

American Association on Intellectual and Developmental Disabilities. (n.d.). *Definition of intellectual disability.* Retrieved January 29, 2019, from http://aaidd.org/intellectual-disability/definition

Battista, M. T. (1983). A complete model for operations on integers. *Arithmetic Teacher, 30*(9), 26–31.

Bouck, E. C., Bassette, L., Shurr, J., Park, J., Kerr, J., & Whorley, A. (2017). Teaching equivalent fractions to secondary students with disabilities via the virtual- representational-abstract instructional sequence. *Journal of Special Education Technology, 32,* 220–231.

Bouck, E. C., Chamberlain, C., & Park, J. (2017). Concrete and app-based manipulatives to support students with disabilities with subtraction. *Educa-*

tion and Training in Autism and Developmental Disabilities, 52, 317–331.

Bouck, E. C., & Park, J. (2018). Mathematics manipulatives to support students with disabilities: A systematic review of the literature. *Education and Treatment of Children, 41,* 65–106.

Bouck, E. C., Park, J., Levy, K., Cwiakala, K., & Whorley, A. (2019). App-based manipulatives and explicit instruction to support division with remainders. *Exceptionality* [Advanced Online Publication]. doi:10.1080/09362835.2019.1586709

Bouck, E. C., Park, J., Shurr, J., Bassette, L., & Whorley, A. (2018). Using the virtual-representational-abstract approach to support students with intellectual disability in mathematics. *Focus on Autism and Developmental Disabilities, 33,* 237–248.

Bouck, E. C., Park, J., Sprick, J., Shurr, J., Bassette, L., & Whorley, A. (2017). Using the virtual-abstract instructional sequence to teach addition of fractions. *Research in Developmental Disabilities, 70,* 163–174.

Bouck, E. C., Satsangi, R., Doughty, T. T., & Courtney, W. T. (2014). Virtual and concrete manipulatives: A comparison of approaches for solving mathematics problems for students with autism spectrum disorder. *Journal of Autism and Developmental Disabilities, 44,* 180–193.

Bouck, E. C., Satsangi, R., & Park, J. (2018). The concrete-representational-abstract approach for students with learning disabilities: An evidence-based practice synthesis. *Remedial and Special Education, 39,* 211–228.

Bouck, E. C. Shurr, J., Bassette, L., Park, J., Kerr, J., & Whorley, A. (2018). Adding it up: comparing concrete and app-based manipulatives to support students with disabilities with adding fractions. *Journal of Special Education Technology, 33,* 194–206.

Bouck, E. C., Working, C., & Bone, E. (2018). Manipulative apps to support students with disabilities in mathematics. *Intervention in School & Clinic, 53,* 177–182.

Brainingcamp. (2018). *Two-color counter.* Retrieved from, https://itunes.apple.com/us/app/two-color-counters/id1052889023?mt=8

Browder, D. M., Spooner, F., Ahlgrim-Delzell, L., Harris, A. A., & Wakemanxya, S. (2008). A meta-analysis on teaching mathematics to students with significant cognitive disabilities. *Exceptional Children, 74,* 407–432. doi:10.1177/001440290807400401

Collins, B. C. (2012). *Systematic instruction for students with moderate and severe disabilities.* Paul H Brookes Publishing.

Doabler, C. T., & Fien, H. (2013). Explicit mathematics instruction: What teachers can do for teaching students with mathematics difficulties. *Intervention in School and Clinic, 48,* 276–285. doi: 10.1177/1053451212473151

Gagnon, J. C., & Maccini, P. (2001). Preparing students with disabilities for algebra. *Teaching Exceptional Children, 34*(1), 8–15.

Gargiulo, R. (2012). *Special education in contemporary society* (4th ed.). Thousand Oaks, CA: Sage.

Gast, D. L., & Spriggs, A. D. (2010). Visual analysis of graphic data. In D. L. Gast (Ed.), *Single subject research methodology in behavioral sciences* (pp. 199–233). Routledge.

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology, 109,* 680–693. doi:10.1037/edu0000159

Ginsburg, H., & Baroody, A. (2003). *TEMA-3 Examiners Manual* (3rd ed.). PRO-ED.

Hart Barnett, J. E., & Cleary, S. (2015). Review of evidence-based mathematics interventions for students with autism spectrum disorders. *Education and Training in Autism and Developmental Disabilities, 50,* 172–185.

Hudson, M. E., Rivera, C. J., & Grady, M. M. (2018). Research on mathematics instruction with students with significant cognitive disabilities: Has anything changed? *Research and Practice for Persons with Severe Disabilities, 43,* 38–53. doi:1.o0r.g1/107.171/1775/4105470976996198187755666O1

King, S. A., Lemons, C. J., & Davidson, K. A. (2016). Math interventions for students with autism spectrum disorder: A best-evidence synthesis. *Exceptional Children, 82,* 443–462. doi:10.1177/0014402915625066

Lord, C., Luyster, R. J., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part II): Toddler Module.* Western Psychological Services.

Maccini, P., & Hughes, C. A. (2000). Effects of a problem-solving strategy on the introductory algebra performance of secondary students with learning disabilities. *Learning Disabilities Research & Practice, 15,* 10–21.

Maccini, P., & Ruhl, K. L. (2000). Effects of a graduated sequence on the algebraic subtraction of integers by secondary students with learning disabilities. *Education and Treatment of Children, 23,* 465–489.

Polloway, E. A., Lubin, J., Smith, J. D., & Patton, J. R. (2010). Mild intellectual disabilities: Legacies and trends in concepts and educational practices. *Education and Training in Autism and Developmental Disabilities, 45,* 54–68.

Powell, S. R., Fuchs, L. S., & Fuchs, D. (2013). Reaching the mountaintop: Addressing the common core standards in mathematics for students with mathematics difficulties. *Learning Disabilities Research & Practice, 28,* 38–48.

Riccomini, P. J., Morano, S., & Hughes, C. A.

(2017). Big ideas in special education: Specially designed instruction, high-leverage practices, explicit instruction, and intensive instruction. *Teaching Exceptional Children, 50*(1), 20–27. doi: 10.1177/0040059917724412

Root, J. R., Browder, D. M., Saunders, A. F., & Yo, Y. (2017). Schema-based instruction with concrete and virtual manipulatives to teach problem solving to students with autism. *Remedial and Special Education, 38,* 42–52. doi:10.1177/0741932516643592

Satsangi, R., & Bouck, E. C. (2015). Using virtual manipulative instruction to teach the concepts of area and perimeter to secondary students with learning disabilities. *Learning Disability Quarterly, 38,* 174–186. doi:10.1177/0731948714550101

Satsangi, R., Hammer, R., & Hogan, C. D. (2018). Studying virtual manipulatives paired with explicit instruction to teach algebraic equations to students with learning disabilities. *Learning Disability Quarterly* [Advanced Online Publication]. doi:10.1177/0731948718769248

Shepley, C., Lane, J. D., & Ault, M. J. (2018). A review and critical examination of the system of least prompts. *Remedial and Special Education* [Advanced Online Publication]. doi:10.1177/0741932517751213

Shurr, J., Jimenez, B., & Bouck, E. C. (2019). *Educating students with intellectual disability and autism spectrum disorder: Book 1 Research-based practices and education science.* Council for Exceptional Children.

Snell, M. E., & Brown, F. (2011). *Instruction of students with severe disabilities* (7th ed.). Pearson.

Spooner, F., Knight, V. F., Browder, D. M., & Smith, B. R. (2012). Evidence-based practice or teaching academics to students with severe developmental disabilities. *Remedial and Special Education, 33,* 374–387. doi:10.1177/0741932511421634

Spooner, F., Root, J., Saunders, A. F., & Browder, D. M. (2018). An updated evidence-based practice review on teaching mathematics to students with moderate and severe developmental disabilities. *Remedial and Special Education.* [Advance Online Publication]. doi:10.1177/0741932517751055

Stephan, M., & Akyuz, D. (2012). A proposed instructional theory for integer addition and subtraction. *Journal for Research in Mathematics Education, 43,* 428–464. doi:10.5951/jresematheeduc.43.4.0428

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development, 93,* 403–411. doi: 10.1002/jcad.12038

Vannest, K. J., Parker, R. I., Gonen, O., & Adiguzel, T. (2016). Single case research: Web-based calculator for SCR analysis. (Version 2.0) [Web-based application]. Texas A&M University. Retrieved July 23, 2018, from singlecaseresearch.org

White, O. R., & Haring, N. G. (1980). *Exceptional teaching* (2nd ed.). Merrill.

Received: 13 February 2019
Initial Acceptance: 9 April 2019
Final Acceptance: 15 May 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 173–184

An Examination of Measurement Comparability for a School Accountability Test among Accommodated and Non-Accommodated Students with Autism

Sara E. Witmer and Sarina Roschmann
Michigan State University

Abstract: Although it is critical for students with autism to be included in large-scale assessment and accountability systems, it is not clear how to best measure their underlying academic skills and knowledge. Additional empirically-supported guidance is necessary to assist school teams that need to make decisions about how to best include students with autism in such systems. Test accommodations may be needed to ensure appropriate measurement on large-scale tests, but it is important to empirically explore whether they are meeting this intended purpose. In the current study, existing statewide test data were analyzed to determine the extent to which measurement comparability was apparent for students with autism receiving and not receiving accommodations on a math test used for accountability purposes. Limited differential item functioning (DIF) was identified for all comparisons, suggesting that existing school-based decisions to include students with autism in regular accountability testing do not correspond to serious measurement concerns. Implications for research and practice are provided.

Over the past few decades, awareness of a need to appropriately include all students, including students with disabilities, in school accountability testing has increased substantially. Federal laws and regulations have required the inclusion of all students in accountability testing, with appropriate accommodation as necessary (ESSA, 2015; IDEA, 2004). These rules are intended to promote access to quality instruction according to the same standards, no matter what the disability status of the child (National Research Council, 1999). Associated regulations indicate that only a very small proportion of all students should be considered proficient via alternate assessments, which may be substantially easier and may lack appropriate alignment to grade-level standards (Elliott et al., 2008). Indeed, some have argued that efforts to include more students in regular testing for accountability purposes have been associated with improved instruction and learning for students with disabilities (Schulte et al., 2001; Ysseldyke et al., 2004). With careful attention to the unique difficulties of students with autism with respect to the Common Core State Standards that they may be expected to meet, and application of associated interventions, it is anticipated that many students with autism will be able to excel across academic content areas (Marsh, 2015).

However, many questions remain about how to best measure the academic knowledge and skills of students with disabilities using existing large-scale assessment programs (Lovett & Lewandowski, 2015). Students with disabilities often have unique characteristics that may make it difficult to appropriately measure their underlying skills and knowledge, particularly when using a common test that is intended to be administered similarly to all students. Altering the typical test conditions

The authors acknowledge the contributions of staff from the Michigan Department of Education in providing data for the analyses included in this article. This research result used data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). Results, information and opinions solely represent the analysis, information and opinions of the author(s) and are not endorsed by, or reflect the views or positions of, grantors, MDE and CEPI or any employee thereof. Correspondence concerning this article should be addressed to Sara E. Witmer, 620 Farm Lane, Rm. 434, Michigan State University, East Lansing, MI 48824. E-mail: switmer@msu.edu

(e.g., providing extended time or reading printed test directions and content aloud), is often pointed to as a potential solution to related concerns. However, at times such changes in test conditions may ultimately reduce the accuracy in measurement of a targeted test construct (Lovett & Lewandowski, 2015). Without further empirical exploration to better understand the conditions under which a student may be appropriately and accurately tested, school teams are left with limited empirical information to guide important decisions about how to include students in large-scale assessment systems. Empirical investigation that is informed by knowledge of the unique characteristics of various students with disabilities, as well as consideration of the conditions under which they may be able to best demonstrate their academic knowledge is needed to help inform decision-making. This is particularly needed for students with autism, who have thus far been neglected in most of the test accommodation research.

Potential Measurement Concerns in Academic Testing for Students with Autism

When considering common characteristics of students with autism in light of typical accountability testing procedures, several areas of possible measurement concern are evident. First, the communication difficulties that students with autism, by definition, experience (American Psychiatric Association, 2013) may reduce their ability to understand test directions or comprehend item content. Relatedly, students with autism may have unique reading comprehension difficulties (Wei et al., 2015), which may affect their ability to understand written test directions or test item content. Second, the sensory processing difficulties that have been noted among many students with autism (Baker et al., 2008) may make it particularly difficult for them to screen out common distractions within a group testing environment. Third, time processing and management difficulties that have been noted to be common among students with autism (Janeslätt et al., 2010) may present unique difficulties when students are expected to complete a test under timed conditions. Finally, students with autism often display problem behaviors (Totsika et al., 2011); these difficulties may interfere with their ability to engage in the sustained academic behaviors needed to obtain an accurate measurement of academic skills through standardized testing. Altogether, these concerns about the potential for inaccurate measurement of academic skills may cause members of individualized education program (IEP) teams to consider alternatives to traditional testing, such as alternate assessment, for students with autism.

Indeed, teacher surveys point to a general tendency to exclude students with autism from regular accountability testing. Using data from the National Longitudinal Transition Study – 2 (NLTS-2) which involved the collection of teacher survey data for a nationally-representative group of high school students with autism in the mid-2000s, Bouck (2017) found that most students were tested using an alternate assessment and very few participated in testing with accommodations. Similarly, through a survey of a statewide sample of teachers of students with autism conducted in 2009-10, Witmer and Ferreri (2014) found that only approximately 30% of students with autism participated in the regular state accountability test. Such widespread exclusion from testing begs the question of whether students with autism are being appropriately included and therefore also whether they have appropriate opportunities to experience the intended benefits of accountability, namely improved instruction and learning.

Early work on the widespread exclusion of students with disabilities from accountability testing indicated the following reasons for their widespread exclusion: students with disabilities were not being instructed according to the standards tested, concern existed about their ability to cope with the anxiety that testing might produce, concern existed that their inclusion would systematically decrease aggregated school and district scores, and there was an unwillingness to offer test accommodations that might facilitate greater participation (Vanderwood et al., 1998). Following the development of test accommodation policies that permitted more liberal use of a variety of accommodations, regular test participation has been found to rise, with less need to provide alternate assessments to many students (Lazarus et al., 2011; Olson & Goldstein, 1996). These findings suggest that

consideration of additional test accommodations targeted to the aforementioned needs of students with autism could help improve test participation. However, further empirical investigation is necessary to ensure that accommodations truly serve to improve the accuracy of measurement among students who are deemed in need of them, instead of merely improving student participation rates.

Testing Accommodations and Measurement Comparability

Test accommodations are changes in test materials or procedures that are intended to remove construct-irrelevant variance that is considered otherwise present due to a student's unique characteristics (Fuchs et al., 2000; National Center on Educational Outcomes, 2018). Over the past few years much empirical work has accumulated to help identify whether specific test changes have their intended accommodating effect for students with disabilities. Studies have indicated that under certain circumstances, certain test changes thought to be accommodations do positively improve the scores of students with disabilities (Buzick & Stone, 2014; Gregg & Nelson, 2012; Li, 2014); however, in some circumstances they have actually been found to have a negative effect on test scores (Elliott et al., 2001). Ultimately, work on test accommodations has pointed to a need to make decisions individually for students instead of based on their disability category (Kevitt & Elliott, 2003), and to objectively examine whether the selected accommodations have the intended positive impact on an individual student's test scores (Fuchs et al., 2000). It also points to the need to provide professional development to help IEP teams make appropriate test accommodation decisions (Shriner & Destefano, 2003). Analysis of state test accommodation policies also point to a perceived need to provide training to individuals who deliver certain accommodations to ensure that they avoid providing inappropriate help on tests (Lazarus et al., 2013).

Interestingly, much of the empirical research on accommodations has focused on their effectiveness in boosting the scores of those students who are deemed in need of the accommodation, rather than on the actual influence of the accommodation on the accuracy of measurement (Lovett & Lewandowski, 2015). Although a unique boost in test score may indeed suggest that construct-irrelevant variance is being removed, it is critical to recognize the possibility that such a boost could actually correspond to a reduction in the accuracy with which the intended academic knowledge and skills are measured. For example, providing a familiar examiner during testing may ultimately result in a greater test score boost for certain students than others, but without careful examination, it may be the case that what is contributing to that test score boost is the fact that the familiar examiner is inappropriately helping specific students by guiding them to the correct responses. In such a situation, using differential boost alone as an indicator of the appropriateness of an accommodation would be flawed, given the improvement in test score does not represent reduction in construct-irrelevant variance, and instead decreases the extent to which the score represents the student's independent achievement with respect to what is intended to be measured.

Increasingly, experts argue that there is a need for measurement comparability studies to ascertain whether test accommodations are indeed facilitating better measurement of intended test constructs (Lovett & Lewandowski, 2015). One approach for examining measurement comparability across student subgroups involves analysis of differential item functioning (i.e., DIF; Walker, 2011). Using this approach, item-level performance data from academic achievement tests administered to various subgroups of students are analyzed. More specifically, item parameters are estimated and compared for focal and reference groups to determine whether measurement concerns are evident (Walker & Beretvas, 2001). If substantial differences in item parameters are identified between focal and reference groups, this may indicate that the test is not effectively measuring the desired construct across both groups. When applying this approach to explore possible measurement concerns for students with disabilities, it can be helpful to first identify a reference group for which there is not expected to be

unique measurement concerns; often, this is non-accommodated students without disabilities. Then, focal groups of both accommodated and non-accommodated students with disabilities can be created, and DIF analysis can be conducted to identify whether DIF exists for either of these groups when compared to the reference group. If DIF is identified for either the accommodated or non-accommodated group of students with disabilities, this may suggest that the test is not measuring the same construct for the given group compared to the reference group. Ideally, one might be able to conduct an experimental study in which the same students with disabilities were tested in an experimental fashion under both accommodated and non-accommodation conditions to help determine whether the accommodation truly reduced the amount of DIF present (Buzick & Stone, 2011); however, such studies are difficult to conduct given the large sample sizes typically needed for meaningful DIF analyses.

Several research teams have explored DIF for accommodated and non-accommodated groups of students with disabilities to understand whether the various testing conditions have been associated with comparable measurement across students (Bolt & Ysseldyke, 2006, 2008; Cook et al., 2010; Finch & Finch, 2013, 2014; Huynh et al., 2004; Kim et al., 2009). Results from these studies have primarily supported the measurement comparability of accommodated tests (Cook et al., 2010; Huynh et al., 2004; Kim et al., 2009), although some have suggested that accommodated test administrations may not fully address concerns about measurement comparability for students with disabilities (Bolt & Ysseldyke, 2006, 2008; Finch & Finch, 2013, 2014). The students included in the associated studies have primarily been those with learning disabilities; no DIF studies were identified that specifically targeted accommodated and non-accommodated students with autism, who may have particularly unique concerns during standardized testing and consequently require different accommodations for measurement comparability to occur.

Existing Empirical Investigation of Testing Accommodations for Students with Autism

Overall, very little empirical work has been done to examine accommodated test administrations among students with autism. The work that has been done has focused on perceptions of what accommodations would be helpful to students with autism. For example, Simpson et al. (1999) explored teachers' perceptions about which accommodations would be helpful to students with autism; teachers frequently reported some of the following as potentially helpful: familiar person present, extended time, computer administration, oral presentation and response, allow marking of answers in test booklet, offering reinforcers, and test breaks. Sarrett (2018) conducted a survey and focus groups of college students with autism who highlighted some of the following as potentially helpful test accommodations: extended time, distraction free test areas, taking breaks, obtaining clear directions.

One recent study was identified that explored accommodation effects for three students with autism on a formative reading assessment (Jones et al., 2018). In the associated study, the typical test condition was paper-based, whereas the accommodated conditions used flash cards and PowerPoint slides that included an auditory signal and timed breaks between items. The hypothesized format by student effects were not identified; however, reliability was found to increase for certain students under the accommodated conditions, suggesting that use of the accommodated methods may allow for more consistent results for certain students.

The Current Study

The current study was intended to address the lack of existing empirical knowledge on measurement characteristics of accountability test administrations for students with autism, a group of students who exhibit particularly unique communication and behavioral characteristics that may interfere with accurate measurement of academic skills under standardized test procedures. The study involved an examination of the measurement comparability of accountability test administrations

for students with autism. More specifically, the research questions for the current study were:

1. To what extent is DIF present on a math test used for accountability purposes when comparing accommodated students with autism to a reference group of non-accommodated students without disabilities?
2. To what extent is DIF present on a math test used for accountability purposes when comparing non-accommodated students with autism to a reference group of non-accommodated students without disabilities?

Such information is needed to help school professionals know whether existing decision-making about inclusion and accommodation for students with autism corresponds to substantial measurement concerns.

Method

Dataset

Student- and item-level performance data from a statewide accountability math test administered in 2012 were sought for analysis. The analyses were conducted solely on the selected-response items (as opposed to constructed response items, which were also included on the test) for students taking the fourth and seventh grade tests. This included 57 items for the fourth-grade test and 60 items for the seventh-grade test. All items were administered via paper-and-pencil format. The specific standards domains that the fourth-grade math test was intended to measure include the following: Operations and Algebraic Thinking, Number and Operations in Base Ten, Number and Operations-Fractions, Measurement and Data, and Geometry. The specific standards domains that the seventh-grade math test was intended to measure include the following: Ratios and Proportional Relationships, The Number System, Expressions and Equations, Geometry, and Statistics and Probability. Coefficient alphas for the associated tests were .90 and .91 for the fourth and seventh grade tests respectively (Michigan Department of Education, 2012). All students in the state were expected to take the regular math test at their corresponding grade levels, with only a very small proportion of students expected to participate in an alternate assessment.

Group Selection

Group selection and analyses were completed separately by grade level. At each grade level, three student subgroups were created to facilitate comparisons of measurement characteristics: (1) a reference group of non-accommodated students reported as not receiving special education services (Ref; $N = 1000$), (2) all students who were reported as receiving special education services for autism and receiving test accommodations (AU+), and 3) all students who were reported as receiving special education services for autism and not receiving test accommodations (AU−). Students who were identified as having autism but only receiving services through Section 504 of the Rehabilitation Act were not included in the analyses. Measurement characteristics of the test for students in the Ref group were anticipated to represent ideal measurement characteristics, given that the students in the reference group did not have disabilities requiring special education services, and took the test without accommodations. They represented a randomly sampled group from all students who were not receiving special education services, and reported as not receiving test accommodations on the math test. The AU groups served as focal groups in the analyses.

Analysis

Initially, a principal components analysis was conducted using item-level data for each student subgroup to determine if the data met the assumption of unidimensionality to permit further analysis using the procedures described below. More specifically, this analysis involved an examination of the ratio of the eigenvalue associated with the first factor to the second factor to determine whether it met the guideline of 5.0 or higher as set forth by Lord (1980).

Next, an application of item response theory was used to estimate differential item functioning (Embretson & Reise, 2000) for each focal group when compared to the corre-

sponding reference group. More specifically, item parameters were estimated for each group in a given comparison (i.e., 4th grade AU+ vs. 4th grade Ref; 4th grade AU− vs. 4th grade Ref; 7th grade AU+ vs. 7th grade Ref; 7th grade AU− vs. 7th grade Ref) using IRTPRO according to a two-parameter logistic model (2PL) and the Bock-Aitkin method (Bock & Aitkin, 1981). Item parameters were estimated for each item using all items in the test as an anchor for the estimation process. Wald tests were used to identify any significant differences in item parameters, with a cut-off of $p < .01$ considered indicative of DIF.

Finally, in order to estimate the magnitude of DIF for each item identified as displaying significant DIF, an adaptation of Wainer's (1993) approach was used. More specifically, using information on the difficulty and discrimination parameter estimates for each subgroup, an item characteristic curve was calculated for each subgroup according to the following equation:

$$P_i(\theta) = \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}}$$

where P represented the probability that a randomly selected examinee with ability θ would answer item (i) correctly, b_i represented the item difficulty value, and a_i represented the discrimination value. Next, the vertical distance between the focal and reference group was calculated at 100 points along the latent trait distribution, and weighted according to the density of the focal group at the given point. These weighted distances were then summed to provide an estimate of the differences in item characteristic curves (ICCs). The following is the associated equation that was approximated using the procedure described above:

$$T(3) = \int_{-\infty}^{\infty} [P_F(\theta) - P_R(\theta)]^2 dG_F(\theta)$$

where $P_F(\theta)$ is the probability of a focal group examinee answering correctly and $P_R(\theta)$ is the probability of a reference group examinee answering correctly, and $dG_F(\theta)$ represents the density of the proficiency distribution of the focal group. The associated cut-off proposed by Roussos and Stout (1996) which suggests effect sizes of .059 or more should be considered moderate DIF was used to evaluate the associated differences in order to determine whether the magnitude of DIF was of particular concern.

Results

Descriptive Information

Demographic information by student subgroup is provided in Table 1. A total of 567 students with autism participated in the fourth-grade test and 593 students with autism participated in the seventh-grade test. At the fourth-grade level, 42%of students with autism were accommodated, whereas at the seventh-grade level 37% of students with autism were accommodated. The accommodations most frequently provided at the fourth-grade level included reader ($N = 84$), Audio ($N = 57$), multi-day administration ($N = 31$) and other ($N = 117$). The accommodations most frequently provided at the seventh-grade level included audio ($N = 97$), reader ($N = 41$), multi-day administration ($N = 9$) and other ($N = 98$).

At both grade levels, the autism groups included larger ratios of males to females (approx. 9:1), whereas the reference group included more equal distributions of males and females. Across both grade levels, the AU subgroups included a smaller proportion of students from lower socioeconomic backgrounds (40%) compared to the reference groups (46 to 48%). The AU subgroups also tended to include smaller proportions of students from minority racial and ethnic backgrounds.

In terms of overall score, the reference group had the highest average raw score, with the AU− subgroup having the next highest average score at both grade levels, and the AU+ subgroup having the lowest average score at both grade levels.

Unidimensionality Analysis

Results of the analysis of unidimensionality are provided in Table 2. All groups met the criteria for unidimensionality to permit additional DIF analysis.

TABLE 1

Demographic Information for Math Test

	Grade 4			*Grade 7*		
	Ref	*AU+*	*AU−*	*Ref*	*AU+*	*AU−*
N	*1000*	*240*	*327*	*1000*	*217*	*376*
Gender (%)						
Female	50.8	13.8	13.8	52.3	10.1	10.9
Male	49.2	86.3	86.2	47.7	89.9	89.1
Race/Ethnicity (%)						
American Indian or Alaska Native	0.5	0.4	0.9	16.7	3.2	1.3
Black or African-American	18.9	13.3	8.3	17.0	8.3	10.1
Hispanic or Latino	5.2	4.2	2.8	5.1	3.2	2.7
Asian	2.8	0.4	3.1	3.1	0	1.1
White	69.0	80.4	82.9	70.7	82.9	82.2
Other	2.8	1.3	2.1	3.7	2.3	2.7
Economically disadvantaged (%)						
No	51.7	60.0	60.9	54.1	59.0	59.6
Yes	48.3	40.0	39.1	45.9	41.0	40.4
Average raw score (standard deviation)	38.8 (10.1)	32.1 (11.2)	37.6 (11.7)	36.8 (11.4)	29.1 (12.5)	34.1 (12.4)

DIF Analysis

Results of the DIF analysis are provided in Table 3. At the fourth-grade level, just two items (3.5% of all items) displayed DIF for the AU+ group, and just three items (5.3% of all items) displayed DIF for the AU− group. At the seventh-grade level, just one item (1.7% of all items) displayed DIF for the AU+ group, and five items (8.3% of all items) displayed DIF for the AU− group. For no items was DIF found to be moderate or large. Figure 1 provides an example of the focal and reference group ICCs for one of the fourth-grade items that showed DIF.

TABLE 2

Results of Principal Components Analysis for Math Test

	Grade 4			*Grade 7*		
	Ref	*AU±*	*AU−*	*Ref*	*AU±*	*AU−*
1	10.0	10.6	12.5	11.3	12.1	12.5
2	2.0	2.3	1.9	2.0	2.1	2.2
3	1.5	1.9	1.7	1.6	2.0	2.0
4	1.4	1.8	1.6	1.5	1.9	1.7
5	1.3	1.7	1.5	1.3	1.7	1.5
Ratio	5.0	4.6	6.6	5.7	5.8	5.7

Discussion

Test accommodations are intended to allow for comparable measurement of constructs across those who are in need of them and those who are not in need of them. However, the extent to which they allow for comparable measurement in current practice remains uncertain. Students with autism, in particular, may need test accommodations to promote

TABLE 3

DIF Analysis Results for Math Test – Autism Groups

	Grade 4		*Grade 7*	
	AU+	*AU−*	*AU+*	*AU−*
N/% of DIF items	2/3.5	3/5.3	1/1.7	5/8.3
N/% of Moderate to Strong DIF items	0/0	0/0	0/0	0/0

more accurate measurement of their academic skills. In the current study, the measurement comparability of a test for students with autism, including those receiving and not receiving test accommodations was explored using an application of item response theory, namely differential item functioning (DIF) analysis. Only very slight concerns with measurement comparability were identified for both accommodated and non-accommodated groups of students with autism; slightly more items displaying small, but significant DIF were identified for non-accommodated students with autism when compared to the number of such items identified for accommodated students with autism.

Overall, these findings seem to suggest that the math skills of students were likely being measured relatively similarly across those with and without autism, as well as across those with autism receiving and not receiving accommodations. The findings align somewhat well with other measurement comparability studies of students with disabilities that have identified relatively few concerns about different measurement qualities for academic tests administered to students with disabilities with and without accommodations (Cook et al., 2010; Huynh et al., 2004; Kim et al., 2009). It suggests that decisions to exclude students with autism from regular academic testing for accountability purposes (i.e., having such students participate in an alternate assessment) may not be justified if the rationale for exclusion is that the test will not measure their academic skills as well as such skills are measured among those without disabilities.

Using data from the U.S. Department of Education (2018), we further explored whether the number of students who participated in the current test appeared to represent the population of students with autism reported to be receiving special education services in the targeted state at the time of test administration. Based on the federal data for the target state, it was estimated that approximately 2200 students with autism would have likely been served in fourth and seventh grade. That estimation, combined with the data used for this study, suggests that only about half of the students with autism likely participated in the state accountability test in math in each of the target grades. Although it is permissible for students with disabilities to participate in an alternate assessment, the findings of the current study appear to beg the question of whether more students with autism could potentially be included in the regular test without threatening the measurement qualities of the test. Doing so could indeed help ensure that their academic outcomes are considered important when school districts make decisions about resource allocation based on statewide accountability test scores.

Although the findings of the current study do not allow for speculation about the actual effectiveness of accommodations, given that accommodation status was not experimentally manipulated, the findings do seem to suggest that, at the very least, accommodations, as provided on the given test, are not associated with highly problematic measurement qualities. Given that the accommodated autism groups actually had slightly more comparable measurement (i.e., less DIF) than the non-accommodated autism groups, it may indeed be the case that they are serving their intended purpose to remove construct-irrelevant variance. However, it must be emphasized that experimental work on accommodations would be needed to fully support that notion.

Although not the focus of this study, it seems important to comment on the differences in average test scores across subgroups. Students with autism, and particularly those receiving accommodations, tended to perform below the reference group of students without disabilities. This seems to suggest there is much more work to be done to close the achievement gap in performance for many students with autism. Information is accumulating on evidence-based practices for improving a variety of outcomes for students with autism (National Professional Development Center on Autism Spectrum Disorder, 2018), although it ultimately remains questionable whether such practices are implemented with fidelity in educational settings (Stahmer et al., 2014). It seems that both use of evidence-based intervention and other supports during in instruction, such as accommodations, may be needed to close the gap.

It also is important to comment on the relatively low proportion of students receiving

accommodations. Although certainly not all students with disabilities are likely to need accommodations, it seems likely that many students with autism are likely to experience difficulties during testing related to their disability characteristics that are not ultimately what the test is intended to measure. For example, difficulties tuning out distractions is a common challenge for students with autism for which various accommodations may be helpful. Reinforcement is another test accommodation that was frequently reported as potentially helpful to students with autism based on prior teacher perception studies (Simpson et al., 1999). In the dataset examined for the current study, such accommodations would have been noted as an "Other Accommodation", which was relatively infrequently reported as being used (i.e., only approximately 19% of students with autism were reported to receive an "Other Accommodation"). One accommodation that has been noted in prior work as potentially helpful to students with autism (i.e., test breaks) is quite similar to the notion of a Multiple Day accommodation, which was only very rarely reported as provided to students with autism in the current study (i.e., 3% were reported to use it). Overall, given the fact that accommodations did not seem to correspond to major problems with measurement in the current study, it seems as if potentially more students with autism, including those students with autism who either participated in the alternate assessment or those who participated in regular test without accommodations, might be able to better demonstrate their underlying academic knowledge and skill with respect to math if additional accommodations were made available.

Limitations

It is critical to note several limitations within the current study. First of all, we were not able to collect data on the integrity with which accommodation use was recorded in the given dataset. Although audits of state accountability testing practices are undertaken to promote quality testing and associated reporting practices, we cannot be entirely sure of the accuracy of accommodation reporting. All students with autism did, in fact, have either a "0" or "1" noted for each accommodation variable, which offers support for the notion that the test administrators did in fact report on accommodation status for each student. However, the overall accuracy cannot be determined. Furthermore, as noted earlier, the study itself did not involve actual experimental manipulation of accommodation status, and so it is not possible to know the actual overall effect of accommodations on the measurement qualities of the test; we only have reason to believe that their provision did not correspond to particularly poor measurement qualities compared to other students. Another potential issue is that student subgroups did vary on demographic characteristics such as gender, race, and sociometric status as opposed to just disability and accommodation status. These may have contributed to measurement characteristic differences. However, it is important to recognize that work was done by the test developers (e.g., review panels, analysis of DIF for gender and race) prior to widespread administration to prevent those particular characteristics from influencing measurement of the intended constructs. In addition, it is important to note that students with autism vary considerably in their characteristics and instructional services received; our study involved grouping them together which hindered our ability to detect differences that may exist for different subgroups of students with autism. Finally, although results of the principal components analysis generally supported the unidimensionality of the construct measured, and therefore the associated DIF analyses were considered appropriate, the unidimensionality results were somewhat borderline, with the accommodation groups, in particular, showing slightly less strong unidimensionality overall.

Implications for Practice and Future Research

Overall, results of this study suggest that many students with autism can be very reasonably included in a large-scale accountability test in math while maintaining reasonable measurement qualities. Given that it appears many such students are still excluded, it seems critical for IEP teams to more critically consider whether they could include more students with autism in regular testing, and consider

use of additional accommodations to potentially promote inclusion. Such work to ensure greater inclusion could help to ensure that the academic needs of students with autism are considered alongside all other students when districts make resource allocation decisions based on accountability test results. Alternatively, if many students with autism continue to be guided toward alternate assessment participation, this may unfortunately contribute to systematically lowered academic expectations and subsequently lower academic performance for these students. Given the limited empirical work on test accommodations that may be particularly helpful for students with autism (e.g., distraction reducers such as noise buffers, various reinforcement tools for item completion, etc.) it seems critical for research to examine the empirical effects of such accommodations, particularly for students with autism who may benefit from them.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Baker, A. E., Z., Lane, A., Angley, M. T., & Young, R. L. (2008). The relationship between sensory processing patterns and behavioural responsiveness in autistic disorder: A pilot study. *Journal of Autism and Developmental Disorders, 38,* 867–875. doi: 10.1007/s10803-007-0459-0

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46,* 443–459. https://doi.org/10.1007/BF02293801

Bouck, E. C. (2017). Understanding participation: Secondary students with autism spectrum disorder and the accountability system. *Education and Training in Autism and Developmental Disabilities, 52,* 132–143.

Bolt, S. E., & Ysseldyke, J. E. (2006). Comparing DIF across math and reading/language arts tests for students receiving a read-aloud accommodation. *Applied Measurement in Education, 19,* 329–355.

Bolt, S. E., & Ysseldyke, J. E (2008). Accommodating students with disabilities in large-scale testing: A comparison of differential item functioning (DIF) identified across disability types. *Journal of Psychoeducational Assessment, 26,* 121–138.

Buzick, H., & Stone, E. (2014). A meta-analysis of research on the read aloud accommodation. *Educational Measurement: Issues and Practice, 33*(3), 17–30. doi: 10.1111/emip.12040

Cook, L., Eignor, D., Sawaki, Y., Steinberg, J., & Cline, F. (2010). Using factor analysis to investigate accommodations used by students with disabilities on an English-language arts assessment. *Applied Measurement in Education, 23,* 187–208. doi: 10.1080/08957341003673831

Elliott, S. N., Kettler, R. J., & Roach, A. T. (2008). Alternate assessments of modified achievement standards: More accessible and less difficult tests to advance assessment practices? *Journal of Disability Policy Studies, 19,* 140–152.

Elliott, S. N., Kratochwill, T. R., & McKevitt, B. C. (2001). Experimental analysis of the effects of testing accommodations on the scores of students with and without disabilities. *Journal of School Psychology, 39,* 3–24. doi: 10.1016/S0022-4405(00)00056-X

Embertson, S., & Reise, S. (2000). *Item response theory for psychologists.* Lawrence Erlbaum Associates.

Every Student Succeeds Act of 2015, Pub. L., No. 114-95, 129 Stat. 1802 (2015).

Finch, W. H., & Finch, M. E. H. (2013). Investigation of specific learning disability and testing accommodations based differential item functioning using a multilevel multidimensional. *Educational and Psychological Measurement, 73,* 973–993. doi: 10.1177/0013164413494776

Finch, W. H., & Finch, M. E. H. (2014). Differential item functioning analysis using a multilevel rasch mixture model: Investigating the impact of disability status and receipt of testing accommodations. *Journal of Applied Measurement, 15,* 133–151.

Fuchs, L., Fuchs, D., Eaton, S., Hamlett, C. L., & Karns, K. (2000). Supplementing teachers' judgments of mathematics test accommodations with objective data sources. *School Psychology Review, 29,* 65–85.

Gregg, N., & Nelson, J. M. (2012). Meta-analysis on the effectiveness of extra time as a test accommodation for transitioning adolescents with learning disabilities: More questions than answers. *Journal of Learning Disabilities, 45,* 128–138. doi: 10.1177/0022219409355484

Huynh, H., Meyer, P. J., & Gallant, D. J. (2004). Comparability of student performance between regular and oral administrations for a high-stakes mathematics test. *Applied Measurement in Education, 17,* 39–57.

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

Janeslätt, G., Granlund, M., Kottorp, A., & Almqvist, L. (2010). Patterns of time processing ability in children with and without developmental disabilities. *Journal of Applied Research in Intellectual Disabilities, 23,* 250–262. doi: 10.1111/j.1468-3148.2009.00528.x

Jones, F., Gifford, D., Yovanoff, P., Otaiba, S., Levy, D., & Allor, J. (2018). Alternate assessment formats for progress monitoring students with intellectual disabilities and below average IQ: An exploratory study. *Focus on Autism and Other Developmental Disabilities,* 1–11. doi: 10.1177/1088357618762749

Kim, D., Schneider, C., & Siskind, T. (2009). Examining equivalence of accommodations on a statewide elementary-level science test. *Applied Measurement in Education, 22,* 144 163. doi: 10.1080/08957340902754619

Lazarus, S. S., Cormier, D. C., & Thurlow, M. L. (2011). States' accommodations policies and development of alternate assessments based on modified achievement standards: A discriminant analysis. *Remedial and Special Education, 32,* 301–308. doi: 10.1177/0741932510362214

Lazarus, S. S., Thurlow, M. L., & Kincaid, A. (2013). *2012 state policies for accommodations used to deliver assessments orally* (Synthesis Report 91). University of Minnesota, National Center on Educational Outcomes.

Li, H. (2014). The effects of read-aloud accommodations for students with and without disabilities: A meta-analysis. *Educational Measurement: Issues and Practice, 33*(3), 3–16. doi: 10.1111/emip.12027

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Lawrence Erlbaum Associates, Inc.

Lovett, B., & Lewandowski, L. (2015). *Testing accommodations for students with disabilities: Research-based practice.* American Psychological Association.

Marsh, W. W. (2015). Common core and the uncommon learner: How autism affects acquisition of common core state standards. *Contemporary School Psychology, 19,* 66–76. doi: 10.1007/s40688-015-0048-8

McKevitt, B. C., & Elliott, S. N. (2003). Effects and perceived consequences of using read-aloud and teacher-recommended testing accommodations on a reading achievement test. *School Psychology Review, 32,* 583–600.

Michigan Department of Education (2012). Michigan Educational Assessment Program Technical Report. Retrieved from: https://www.michigan.gov/documents/mde/MEAP_2010-2011_Technical_Report_394693_7.pdf

National Center on Educational Outcomes. (2018, June 27). *Frequently Asked Questions* [Webpage]. Retrieved from: https://nceo.info/Resources/faq/general_assessments/accessibility_accommodations

National Professional Development Center on Autism Spectrum Disorders. (2018, June 28). *Evidence-based Practices* [Webpage]. Retrieved from https://autismpdc.fpg.unc.edu/evidence-based-practices.

Olson, J. F., & Goldstein, A. A. (1996). Increasing the inclusion of students with disabilities and limited English proficient students in NAEP. *Focus on NAEP, 2*(1). National Center for Education Statistics.

National Research Council. (1999). *Testing, teaching, and learning: A guide for states and school districts.* Committee on Title I Testing and Assessment, R. F. Elmore & R. Rothman (Eds.). National Academy Press.

Roussos, L., & Stout, W. (1996). A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement, 20,* 355–371.

Sarrett, J. C. (2018). Autism and accommodations in higher education: Insights from the autism community. *Journal of Autism and Developmental Disorders, 48,* 679–693. doi: 10.1007/s10803-017-3353-4

Schulte, A., Villwock, D., Whichard, S., & Stallings, C. (2001). High stakes testing and expected progress standards for students with learning disabilities: A five-year study of one district. *School Psychology Review, 30,* 487–506.

Shriner, J. G., & Destefano, L. (2003). Participation and accommodation in state assessment: The role of individualized education programs. *Exceptional Children, 69,* 147–161. doi: 10.1177/001440290306900202

Simpson, R. L., Griswold, D. E., & Myles, B. S. (1999). Educators' assessment accommodation preferences for students with autism. *Focus on Autism and Other Developmental Disabilities, 14,* 212–219, 230. doi: 10.1177/108835769901400403

Stahmer, A. C., Rieth, S., Lee, E., Reisinger, E. M., Mandell, D. S., & Connell, J. E. (2015). Training teachers to use evidence-based practices for autism: Examining procedural implementation fidelity. *Psychology in the Schools, 52,* 181–195. doi: 10.1002/pits.21815

Totsika, V., Hastings, R. P., Emerson, E., Lancaster, G. A., & Berridge, D. M. (2011). A population-based investigation of behavioural and emotional problems and maternal mental health: Associations with autism spectrum disorder and intellectual disability. *Journal of Child Psychology and Psychiatry, 52,* 91–99.

U.S. Department of Education. (2018). IDEA Section 618 data products: State level data file Part B child count and educational environments (2016). Retrieved from: https://www2.ed.gov/programs/osepidea/618-data/state-level-data-files/index.html

Vanderwood, M., McGrew, K., & Ysseldyke, J. E. (1998). Why we can't say much about students with disabilities during educational reform. *Exceptional Children, 64,* 359–370.

Wainer, H. (1993). Model-based standardized measurement of an item's differential impa t. In P.W.

Holland & H. Wainer (Eds.) *Differential item functioning* (pp. 123–135). Lawrence Erlbaum Associates.

Walker, C. M. (2011). What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. *Journal of Psychoeducatoinal Assessment, 29,* 364–376. https://doi.org/10.1177/0734282911406666

Walker, C. M., & Beretvas, S. N. (2001). An empirical investigation demonstrating the Multidimensional DIF paradigm: A cognitive explanation for DIF. *Journal of Educational Measurement, 38,* 147–163.

Wei, X., Christiano, E. R. A., Yu, J. W., Wagner, M., & Spiker, D. (2015). Reading and math achievement profiles and longitudinal growth trajectories of children with an autism spectrum disorder. *Autism, 19,* 200–210. doi: 10.1177/1362361313516549

Witmer, S. E., & Ferreri, S. J. (2014). Alignment of instruction, expectations, and accountability testing for students with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 29,* 131–144. doi: 10.1177/1088357614522294

Ysseldyke, J., Nelson, J. R., Christenson, S., Johnson, D. R., Dennison, A., Triezenberg, H., & Hawes, M. (2004). What we know and need to know about the consequences of high stakes testing for students with disabilities. *Exceptional Children, 71,* 75–94. doi: 10.1177/001440290407100105

Received: 30 January 2019
Initial Acceptance: 28 March 2019
Final Acceptance: 15 July 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 185–200

Using Peer-Mediation to Enhance Conversation and Reduce Inappropriate Communication Acts in Adolescents with Autism

Amanda Thomas and Linda M. Bambara
Lehigh University

Abstract: Pragmatic language difficulties of adolescents with autism spectrum disorder (ASD) can significantly impact their ability to engage in socially appropriate conversations and form peer relationships. Using a multiple baseline design across participants, this study evaluated the effects of a novel peer-mediated intervention (PMI) on improving the social conversation of three high school students with ASD who engaged in high rates of inappropriate communication acts (e.g., perseveration, abrupt topic shifts). The PMI incorporated peer training, graphic/text cues and direct instruction for the students to reduce inappropriate communications while supporting appropriate initiations and topic maintenance responses. Results revealed that the PMI was highly effective in reducing the inappropriate communication acts of three students and increasing the appropriate communication acts of two students. Conversation gains generalized to untrained peers for all student participants. Social validity outcomes assessing intervention acceptability and conversation were highly positive. This study provides additional evidence that PMI can be individualized to address the unique pragmatic language needs of high school students with ASD.

For adolescents with autism spectrum disorder (ASD), pragmatic language, or the use of language in social contexts, is an important concern. In high school, students with ASD encounter a more complex social landscape as social interaction with peers revolves around conversation (Carter et al., 2005). Consequently, with deficits in pragmatic functioning and the increased demand for social communication, high school students with ASD may have difficulty engaging in socially appropriate conversation with their peers, resulting in social isolation and decreased opportunities for peer relationships (Locke et al., 2010).

The pragmatic language difficulties of individuals with ASD are well documented. During conversations, individuals with ASD often display difficulty with both topic management, or appropriately introducing and maintaining topics of conversation, and conversational reciprocity, or engaging in the to-and-fro of conversation (Paul et al., 2009). Common difficulties include the use of irrelevant detail, inappropriate or abrupt topic shifts, topic perseveration, lack of responsiveness, poor reciprocity, and the use of semantically vague referents (Paul et al., 2009).

Pragmatic language difficulties may vary widely from student to student, requiring individualized interventions. Fey's (1986) conversational model provides a useful taxonomy for identifying patterns of pragmatic difficulties and relevant goals for intervention. The model categorizes pragmatic language functioning into four profiles based the degree of assertive acts that are not directly solicited by a communication partner (e.g., introducing new topics, asking questions, adding topic related information) and responsive acts that are contingent upon and responsive to a partner's communications (e.g., responding to questions, acknowledging partner comments). Competent conversationalists, according to Fey, are both assertive and responsive. Of interest to this study are those described by Fey as verbal non-communicators, who are highly assertive and very talkative, but also unresponsive or neglectful of

Correspondence concerning this article should be addressed to Amanda Thomas, Lehigh University, III Research Drive, Bethlehem, PA 18015. E-mail: amd613@lehigh.edu

their partner's communications. These conversationalists often inappropriately dominate interactions by abruptly changing topics, engaging in high rates of contextually unrelated responses, asking many questions on topics of self-interest (sometimes answering their own questions), failing to respond to requests for information, and engaging in incoherent lengthy narratives (Fey, 1986; Fujiki & Brinton, 1991). These characteristics are consistent with the perseverative speech of some individuals with ASD (e.g., Fisher et al., 2013) and capture many of the pragmatic difficulties of this population reported by Paul et al. (2009). Appropriate intervention goals for these communicators include increasing appropriate initiations of new topics, while also increasing responsiveness to partners and the relatedness of their conversational acts to the topic under discussion (Fey, 1986; Paul, 2008).

Social communication intervention research conducted with high school students with ASD is scarce (Kucharczyk et al., 2015; Wong et al., 2015). Peer-mediated intervention (PMI), in which typical peers are taught strategies for improving the social interactions of students with ASD, holds strong promise for addressing the pragmatic communication needs of high school students with ASD. First, it has emerged as an evidence-based practice for addressing a broad range of social communication outcomes for students with ASD (Wong et al., 2015). Second, PMI can be tailored to target specific pragmatic goals, allowing for individualization (Thiemann & Goldstein, 2004). Third, because of peer involvement, PMI fosters improved competence in natural social contexts, an essential goal for any communication intervention (Paul, 2008). Finally, PMI has been shown to successfully promote skill generalization of students with ASD across people and settings (e.g., Hughes et al., 2011; Kamps et al., 2014).

PMI studies specifically focused on improving the reciprocal communication or conversational skills of students with ASD, typically employ a multicomponent intervention consisting of peer and focus student training paired with text or graphic (picture) cues to elicit targeted communication skills (e.g., Kamps et al., 2014, Thiemann & Goldstein, 2001). For example, Kamps et al. (2014) combined peer network training with the direct instruction of four elementary students with ASD and the use of text cues for communication acts related to play interactions (e.g., requests, niceties, comments). Results indicated substantial improvements in the the focus students' total communication acts during free play with trained network peers for all four students and generalization to untrained peers for three of the four students. Further, to optimize communication outcomes, research supports the necessity of a multicomponent PMI approach, as neither peer or focus student training may be sufficient alone (Bambara et al., 2016; Goldstein et al., 2007; Thiemann & Goldstein, 2004). For example, Thiemann and Goldstein (2004) found peer training alone increased general interactions between trained peers and students with ASD, but was inadequate in improving the focus students' use of specific communication skills. Improvements occurred only after peer training was combined with direct focus student instruction on how to use text cues for specific conversational targets (e.g., initiations, comments).

Only a few PMI studies have measured improvements in the conversational abilities of adolescents with ASD in high school settings. Hughes et al. (2000, 2011, 2013) used communication books to encourage interactions between high school students with ASD and typical peers. Students with ASD were taught how to use the communication books and peers were taught to prompt the focus students to use the book and support conversation. Results indicated increased conversational interactions (e.g., initiations and responses) by students with ASD across high school settings. One drawback of the intervention in these studies was the limited structure of the communication book, which contained a finite set of topics or conversational openers for all social interactions. In two studies, Bambara, Cole, et al. (2016) and Bambara et al. (2018) used a multicomponent PMI to increase the assertive conversation skills of high school students with ASD who were passive communicators (e.g., responsive to peers, but rarely assertive). Peers were taught three strategies to (a) support conversational interactions, (b) promote initiations made by the focus student, and (c) increase follow-up questions made by the focus student to sustain

conversation. Additionally, the focus students were taught to use two cue cards with written text to elicit initiations and follow-up questions. The students selected topics for discussion and text cues for initiations changed daily to encourage varied conversations. Results in both studies indicated the students with ASD exhibited higher rates of communication acts, increased initiations and follow up questions, and engaged in longer conversations with trained peers during lunch.

The purpose of this study was to extend PMI research in high school settings by evaluating the effects of a novel PMI on improving the social conversation of students with ASD who can be described as verbal non-communicators. Only a few studies involving young children with ASD have attempted to measure the impact of a PMI on reducing inappropriate communication (e.g., Chung et al., 2007; Thiemann & Goldstein, 2001). Thus, this study involving adolescents with ASD who engaged in high rates of inappropriate talking addresses a critical gap. Specifically, we sought to improve the social conversation of high school students with ASD by reducing inappropriate communication acts that were perseverative, unresponsive, or unrelated to peer verbalizations or the topic of conversation, while increasing the use of appropriate communication acts in the form of initiations and topic maintenance responses. Building upon previous PMI research, this study used a multicomponent intervention consisting of peer training and direct student instruction on how to use graphic cue cards to appropriately initiate and maintain on topic responses. In addition, we assessed whether the students' improved conversations would generalize to untrained peers. We addressed the following questions: (1) What are the effects of the PMI on inappropriate and appropriate communication acts made by high school students with ASD in conversations with trained network peers? (2) To what degree will the PMI result in the generalization of targeted gains (e.g., decreased inappropriate and increased appropriate acts) with untrained peers? (3) How will the students with ASD, trained network peers, and school educators naïve to the intervention, judge the acceptability and outcomes of the PMI?

Method

Participants: Students with ASD and Peers

Focus students. Three focus students with ASD participated. They attended a special education program for students with extensive support needs in a mid-size, diverse urban high school located in the Northeast US. Selection criteria were (a) a primary clinical or educational diagnosis of ASD, (b) an ability to speak in phrases or complete sentences across a range of communicative functions (e.g., request information, comment) and, (c) an expressed interest in participating in the study (i.e., willingness to talk to new friends and learn new ways to have conversations). In addition, we sought students who engaged in frequent inappropriate communication acts (e.g., inappropriately changing topics, reintroducing a previously discussed topic, engaging in non-interactive monologues) during conversations that resulted in ignoring their conversational partner's communication attempts. Teachers first nominated students meeting the criteria, and we then verified eligibility through observations and record review. The Childhood Autism Rating Scale, 2nd Ed (CARS-2; Schopler et al., 2010) and the Social Responsiveness Scale 2 (SRS-2; Constantino & Gruber, 2012) were administered to describe the students' severity of autism symptoms and social behavior deficits.

Josh, a 17-year-old White male, had secondary diagnoses of speech and language impairment, and intellectual disability (KABC-II Nonverbal index of 47). His scores on the CARS-2 (Total raw score = 36) and the SRS-2 (T-score = 81) were both in the severe range for symptoms of ASD and social behavior deficits. Josh was highly talkative, but he often ignored his partners' communication by interrupting to introduce perseverative topics (e.g., *Home Alone*, holidays, scary movies), asking and answering his own questions, and scripting lines from TV shows or movies. Josh had difficulty staying on topic for more than three reciprocal turns, often lapsing into in lengthy monologues that were non-interactive.

Laisha, a 15-year-old Hispanic female, also had a secondary diagnosis of speech and language impairment. Her scores on the CARS-2 (Total raw score = 42) and *SRS-2* (T-score =

90) were both in the severe range for symptoms of ASD and social behavior deficits. Laisha showed interest in maintaining social interaction with her partners by frequently initiating, smiling, and maintaining eye contact. Her verbalizations, including initiations and responses, were frequently vague one- or two-word utterances, making reciprocal interactions difficult. Additionally, during conversations, Laisha often attempted to communicate by scripting lines and listing characters from TV shows and movies in a non-interactive, perseverated dialogue. She often interrupted her partner to introduce a topic of interest to her.

Jerome, a 15-year-old Black male, was also diagnosed with a speech and language Impairment. His scores on the CARS-2 (Total raw score = 44) and SRS-2 (T-score = 73) were in the severe range for symptoms of ASD and moderate range for social behavior deficits. Jerome was often highly assertive or socially avoidant when interacting with peers or adults. When highly assertive, Jerome frequently dominated the interaction by engaging in long, emotional monologues about a time when he was upset or got into trouble. He rarely paused for his partner to comment or ask questions, and often ignored his partner's requests for information. On the other hand, Jerome, at times, avoided conversational interactions by laying his head down on a table or turning his back to peers and adults when they attempted to talk to him.

Peers. Prior to the start of the study, focus students had very little opportunity to interact with general education peers. To recruit peers, we made presentations at service clubs and sought teacher nominations for responsible students. Criteria for inclusion were an expressed interest in interacting with students with ASD and availability. Ten peers, ages 15–17 years old participated. Seven of these students were organized into peer networks for training, one network for each student with ASD (Josh, three peers; Laisha, two peers; Jerome, two peers). The remaining three students served as generalization peers and remained untrained throughout the study. The peers (one male and nine females) were racially and ethnically diverse (four Black, five Hispanic, one White).

Setting

All experimental observations of the focus students' conversations with peers took place in the cafeteria during the focus students' regularly scheduled breakfast (Jerome) or lunch (Josh and Laisha). During observations, the focus student sat at a table and ate with two network or generalization peers. No other students were present at the table and no other activity was introduced. Josh's three network peers took turns rotating across observation sessions depending on their availability. Peer and focus student training took place in an empty classroom during the breakfast or lunch period when no observations were conducted.

Focus Student Conversation and Peer Cue Cards

Conversation card. During intervention, focus students were provided a conversation card to help them appropriately initiate and maintain on-topic responses. The 8 × 4-in card depicted five conversational topics. Each topic was represented by two thematically-related 1 × 1.5-in pictures laid side by side (or in a row). The first picture represented a general topic for discussion such as a favorite TV show (e.g., SpongeBob SquarePants) or video games. The second picture, which was related to the first, provided a cue to talk more about the topic such as SpongeBob favorite characters or episodes, or a Pokémon video game. In addition, Laisha's and Jerome's cards included written phrases in the focus student's own words above the pictures (e.g., "My favorite show is . . .") to cue appropriate verbalizations.

Conversation topics or pictures changed daily and incorporated the focus student's interests including school or home activities, future events, and hobbies. Prior to each intervention session, the focus student and the instructor generated ideas for conversation topics (e.g., "What would you like to talk to your friends about?), and then selected a representative image for each topic from the internet. The instructor encouraged focus students to select varied topics within their areas of interest. To create the second picture to cue talking more about a topic, the instructor asked, "What else do you want to tell them

about _____?" Once the second image for each topic was selected, all pictures were printed to create the conversation card with additional written text for Laisha and Jerome.

Peer cue cards. Network peers were also provided with a cue card during intervention. The 3 × 5-in card summarized the five conversation strategies taught to the peers to facilitate appropriate focus student initiations and topic maintenance responses (see Peer Training).

Dependent Measures and Data Collection

Using a small digital camera and wireless microphone, 10-min samples of the focus students' breakfast or lunch conversations with peers were video recorded. Observations occurred 3–4 days per week over 4 months and across experimental conditions. Once all participants were seated at the table, recording began. Adults moved from the participants' view and did not interfere with their interaction. We used 10-s partial interval recording system to collect data on the percentage of intervals with appropriate and inappropriate communication acts made by the focus student. After each session, a trained observer viewed the video marked with audible 10-s beeps and recorded the occurrence of each communication act in each interval on a data sheet. Appropriate and inappropriate acts could be scored in the same interval and were segmented by a conversational turn or change in communication act as defined below. A communication act is defined as any verbalization (e.g., statement or question) or gesture (e.g., head nod, pointing to the conversation card) directed to a conversational partner. Observers viewed the video as many times as necessary to accurately code, making corrections as needed.

Adapting codes from Bambara, Cole, et al. (2016) and Chung et al. (2007), an appropriate communication act was defined as an initiation (i.e., introduces a new topic unrelated to the prior communication act) or a topic maintenance response (i.e., is related to the topic of conversation and/or peers' immediate prior communication act) that was directed to the peers and semantically clear. Inappropriate communication acts included (a) abrupt topic changes, (b) semantically vague utterances, (c) perseverative speech, (d) no response or refusals, and (e) other non-communicative verbalizations or actions (see Table 1 for definitions).

Inter-observer Agreement (IOA). Prior to the start of the study, a pool of six observers who remained naïve to the study's procedures were trained to 80% criteria on the dependent measures. On at least 30% of the observation sessions, randomly selected for each student and in each experimental condition, two independent observers coded the same 10-min video to establish point-by-point IOA (Kazdin, 2011). Mean agreement for appropriate communicative acts was 85% (range = 78–93%) for Josh, 92% (range = 80–100%) for Laisha, and 93% (range = 81–100%) for Jerome. Mean agreement for inappropriate communicative acts was 83% (range = 75–100%) for Josh, 95% (range = 77–100%) for Laisha, and 92% (range = 75–100%) for Jerome.

Experimental Design and Conditions

A multiple baseline design across participants (Kazdin, 2011) was used to evaluate the effects of the PMI on the focus students' conversation with their peers. Experimental conditions were (a) baseline, (b) peer and focus student training, (c) the peer mediation intervention, and (d) generalization across untrained peers. No data were collected during peer and focus student training. After a minimum of five baseline observations and a predictable pattern of responding, peer-mediated training and intervention was sequentially introduced to each focus student allowing for the demonstration of experimental control at three different points in time. The first author was the instructor for peer and focus student trainings.

Procedure

Baseline. Prior to baseline, network peers and the focus student were introduced and ate together for several days to become acquainted with one another and accustomed to the video camera. During pre-baseline and baseline, the focus student and peers were instructed to "Be friendly" and "Interact like you would with your friends." No further instruction was provided.

TABLE 1

Operational Definitions of Social Communication Behaviors

Communication Act	*Definition*
Appropriate	A communication act that functions as (a) An **initiation** that introduces a new topic (i.e., is unrelated to the prior communication act) or (b) a **topic maintenance response** that is related to the ongoing topic of conversation and/or a peer's immediate prior communication act. To be scored as appropriate, the conversational act must be clearly directed to a peer and be semantically clear. Appropriate acts include verbal utterances and the following nonverbal gestures: pointing to a topic card, and head nod in response to peer questions.
Inappropriate	
Topic Change	A communication act that introduces a new topic (i.e., initiates) before responding to a peer request for information (e.g., P: "What did you do this weekend?" FS: "What is your favorite character in Scooby Doo?") or interrupts (definite overlap of words) a peer to introduce a topic.
Semantically Vague	A communication act seemingly unrelated to the topic of conversation that is vague or unclear in its meaning (i.e., semantic referent is unrecognizable). Example: P: "What did you do this weekend?" FS: "Rainbow Dash."
Perseverative Speech	A communication act that (a) reintroduces a topic that has been previously discussed during the session, (b) repeats a question or requests information more than once (does not include repeating for clarification), or (c) asks a question and then immediately answers it (e.g., "Who is Marlin? Nemo's dad.").
No Response/Refusals	The FS does not respond within 5 s to a peer request. Includes saying "I don't know" when used to avoid responding.
Other	A non-communicative or non-interactive act such as (a) noises, (b) self-talk or scripting (e.g., stating lines from movies), (c) monologue consisting of a continuous string of utterances that lasts longer than two consecutive 10 s intervals and prohibits peers from interacting (e.g., FS continues talking over peer), or (d) inappropriate orientation (e.g., head down on the table).

Note: FS = Focus Student; P = Peer.

Peer training. Peer networks met with the instructor for five 20-min training sessions. Peers were taught five conversation strategies to help the focus student appropriately (a) initiate new topics of conversation using the conversation card, and (b) maintain on-topic responses following either a focus student or peer initiation. During each session, the instructor implemented one or more of the following training components: (a) a rationale for the goals of the PMI intervention or a particular strategy; (b) written examples of strategies; (c) a demonstration of strategy use; (d) opportunities for peers to role-play and practice strategies; and (e) feedback on peer performance during role-play.

In the first training session, the instructor (a) introduced the goals of the PMI, and (b) overviewed the conversation strategies to support appropriate initiations and on-topic responses. The instructor and peers also discussed general characteristics of autism and the individual characteristics of the focus student, identifying appropriate and inappropriate conversation acts.

In the second session, the instructor taught peers the first strategy on how to create opportunities for, and if necessary prompt the

focus student to appropriately initiate a topic of conversation using his or her conversation card as needed. Specifically, peers were taught to pause, look expectantly, and wait up to 10 s for the focus student to initiate. If the focus student did not initiate after a pause, or after being directed to talk about something else (see Move/On Redirect) peers were taught to prompt focus student to use the conversation card by pointing to the card and/or saying, "What do you want to talk about?"

In the third session, the instructor taught peers four facilitative strategies (Bambara et al., 2016) to help the focus student appropriately maintain on-topic responses following an initiation. The strategies were: (a) *Show Interest* (i.e., as your friend talks smile/nod, respond to questions, make positive statements about the topic [e.g., "Cool"]), (b) *Keep the Conversation Going* (i.e., extend the topic for more than three turns by asking questions about it, sharing information on the topic), (c) *Help Your Friend Respond* (i.e., call your friend's name if distracted, ask again or rephrase a question if your friend does not respond, ask for clarification if you do not understand what your friend said), and (d) *Redirect/Move On* (i.e., redirect back to the conversation if your friend changes the topic before you both talked about it [e.g., "Let's finish talking about this first"] or move on or encourage a topic change if your friend repeats a topic [e.g., "We already talked about this before." "Let's talk about something different."). In the fourth session, peers practiced combining all strategies. With the instructor playing the role of the focus student, peers practiced (a) pausing for and prompting appropriate initiations, and (b) using the facilitative strategies to encourage topic maintenance responses around a single topic for as long as there was interest.

In the fifth session, the focus student joined the peers. The instructor first modeled conversation and use of the strategies with the focus student using the conversation card. Then, each peer practiced with the focus student for 5 min with instructor providing verbal direction as needed. Once the conversation ended, the instructor gave feedback to both the peer and focus student. The session ended when all peers demonstrated using the strategies to promote appropriate initiations and topic maintenance responses with the focus student.

Focus student training. Concurrent with peer training, the focus student attended four 30 min training sessions, receiving direct instruction on how to use the conversation card. Every session followed the same format. First, the instructor gave a rationale for having "good" conversations. Second, the instructor explained and demonstrated how to use the card. The instructor modeled how to initiate by making a statement or asking a question about the topic in the first picture and then modeled continuing the conversation by making a statement or asking a question using the second picture. Third, the focus student then practiced initiating and making a topic related response as the instructor pointed to the pictures. Finally, the instructor role-played as the peer, using the conversation strategies to give the focus student opportunities to practice initiating and engaging in on-topic responses for all topics on the card.

Peer-mediated intervention. After training, the peers and focus student were directed to use their strategies during their breakfast or lunch conversations. Prior to each intervention session, the instructor continued to meet with the focus student to create the conversation card. Once in the cafeteria, the instructor gave the focus student his/her card and briefly checked in with the peers, asking if they had any concerns or questions. Peers were also given their cue cards as a reminder to use their strategies as needed to encourage initiations and facilitate on-topic conversations. Once peers joined the focus student at the table, the session started and there was no further instructor interaction.

Generalization. Generalization probes with two untrained peers were conducted approximately after every sixth observation session across experimental conditions. Baseline procedures were followed. Generalization peers and focus students were introduced and ate together prior to data collection to become acquainted and they were instructed to be friendly and interact like they would with their friends. No additional instruction was given. The focus student was provided with a conversation card on probe sessions conducted after training.

Training and Procedural Fidelity

Trained observers measured training fidelity on 100% of all peer and 89% of all focus student training sessions using a fidelity checklist to indicate whether core training components (e.g., rationale, modeling, role-play, feedback) were implemented as intended. Checklist items ranged from 3–6 depending on session content. Mean training fidelity was 97% (range = 91–100%) for peers, and 97% (range = 90–100%) focus students. Procedural intervention fidelity was also assessed by observers using a three-item checklist that measured the instructor checking in with the peers and having all materials present during the intervention. Mean fidelity was 100% and assessed on 92% of all intervention sessions distributed across focus students.

Strategy Use

Trained observers also coded videos to assess peers' and focus students' use of the PMI strategies during intervention. At the end of each 1-min segment, observers used a checklist to record the occurrence of each of the five conversation strategies used by any network peer to promote appropriate conversation as defined (see Peer Training). Observers also recorded, within each 1-min segment, the focus student's use of the conversation card to initiate and/or maintain conversation (i.e., the focus student clearly referred to card and made a statement or asked a question about the topic as an initiation or topic maintenance response). Scoring yielded a mean percentage of intervals per session in which each strategy was implemented by the peers or focus student. IOA for strategy use, calculated point-by-point, was conducted on 33% of the observation sessions for each focus student and peer network. Mean agreement for peers' and focus students' use of each strategy was 80% or higher. Overall mean agreement across all strategies was 96% (80-100%) for peers and 95% (80-100%) for focus students.

Social Validity

Social validity was assessed in three ways. First, upon conclusion of the study, network peers completed an 11-item survey adapted from the *School Intervention Rating Form* (*SIRF*: Harrison et al., 2016). Peers rated the ease of implementing and learning the peer-mediated strategies, the acceptability of the peer training and peer-mediated intervention strategies, the perceived benefits of the intervention for both the focus student and themselves, and the perceived effectiveness of the peer-mediated intervention using a 7-point Likert-type scale ranging from 1 = strongly disagree to 7 = strongly agree. Second, also at the end of the study, focus students completed a 5-item survey, rating their satisfaction, enjoyment, and ease of having conversations with their peers using a 5-point Likert-type scale ranging from 1 = not at all or never to 5 = a lot, very easy, or always. Third, two administrators, two special education teachers, and two paraprofessionals from the high school who were blinded to the intervention rated changes in the quality of the social conversations between the focus students and peers. Two 3-min representative video samples selected from the last three sessions of baseline and intervention were presented to the educators with the order of the focus students and experimental conditions counterbalanced for each showing. After viewing the baseline and post training video, the educators independently rated the conversational behaviors of the focus student, the peers, and the overall quality of the social conversation using a 5-point Likert-type scale ranging from 1 = strongly disagree, 3 = neutral, 5 = strongly agree. Items are shown in Table 2.

Data Analysis

Visual analysis (e.g., changes in level, trend, variability, and immediacy of effect) was first conducted to inspect changes across experimental conditions for the graphed dependent measures. Next, using a web-based calculator (Vannest et al., 2016), Tau-U, a non-parametric effect size index that considers the magnitude of change from baseline to intervention, was calculated for each dependent measure. The simple form of Tau-U (i.e., without trend control) was used to calculate the percentage of all non-overlapping data minus all overlapping data, yielding a score from 0 to 1.0. Tau-U effect sizes from 0–.65, .66-92, and .93–1.0, are inter-

TABLE 2

Summary of Educator Ratings of the Quality of Conversational Interactions During Baseline and Intervention Conditions

	Josh		*Laisha*		*Jerome*	
	B	*I*	*B*	*I*	*B*	*I*
Focus student interactions						
1. FS is responsive to peers' questions or comments.	3.50 (3-5)	4.50 (4-5)	3.50 (1-4)	4.67 (3-5)	3.17 (1-5)	4.75 (4-5)
2. FS is engaged in appropriate conversational turn-taking with peers.	3.17 (2-4)	4.33 (4-5)	3.33 (1-5)	4.67 (4-5)	2.67 (1-5)	4.75 (4-5)
3. FS shows interest in the peers and what peers are saying.	3.67 (2-4)	4.67 (4-5)	3.00 (1-4)	4.67 (4-5)	2.83 (1-5)	4.50 (4-5)
Peer interactions						
4. Peers attend to and attempt to facilitate appropriate on-topic conversation with FS.	4.33 (3-5)	4.83 (4-5)	4.17 (4-5)	4.83 (4-5)	3.83 (3-5)	4.50 (4-5)
5. Peers encourage sustained conversation interaction	4.33 (3-5)	5.00 (5)	3.83 (2-5)	4.50 (4-5)	3.50 (3-5)	4.83 (4-5)
6. Peers show interest in what the FS is saying	4.17 (3-5)	4.83 (4-5)	3.83 (3-5)	4.67 (4-5)	3.50 (2-4)	4.17 (3-5)
Overall Conversation						
7. Conversation is comfortable and natural	3.17 (2-4)	4.33 (4-5)	3.33 (1-4)	4.67 (4-5)	2.17 (1-4)	4.17 (4-5)

Note: FS = Focus Student.

preted as minimal to no effect, moderate effect, and large effect, respectively.

Results

Inappropriate vs Appropriate Communicative Acts

Figure 1 shows the percentage of intervals with inappropriate and appropriate communication acts for the focus students. Overall, following training and the onset of PMI, all three focus students demonstrated a substantial decrease in inappropriate acts, while appropriate communication acts increased or became less variable relative to baseline. Further, a clear separation of data paths for appropriate and inappropriate communication was evident during PMI for all participants.

During baseline, Josh exhibited high, stable rates of inappropriate communication acts ($M = 61.5$), while his appropriate communication acts ($M = 56$) were more variable and decreased just prior to training. With the onset of PMI, an immediate change in level and trend occurred for both measures. Josh's inappropriate communication acts substantially decreased ($M = 18.3$), while his appropriate acts increased above baseline rates ($M = 86$). As shown in Table 3, Tau-U indicated a large intervention effect for both inappropriate (1.0) and for appropriate (1.0) communication acts. His participation ended when he changed school districts.

During baseline, Laisha's inappropriate ($M = 39$) and appropriate ($M = 48$) communication acts were highly variable and overlapped, with slightly higher levels of appropriate than inappropriate talking. With implementation of PMI, her inappropriate communication acts immediately decreased to consistently low rates ($M = 3$), with no inappropriate acts occurring during last five sessions. Appropriate communication acts during PMI were less variable and stabilized at rates slightly higher than baseline ($M = 55$). Tau-U indicated a large intervention

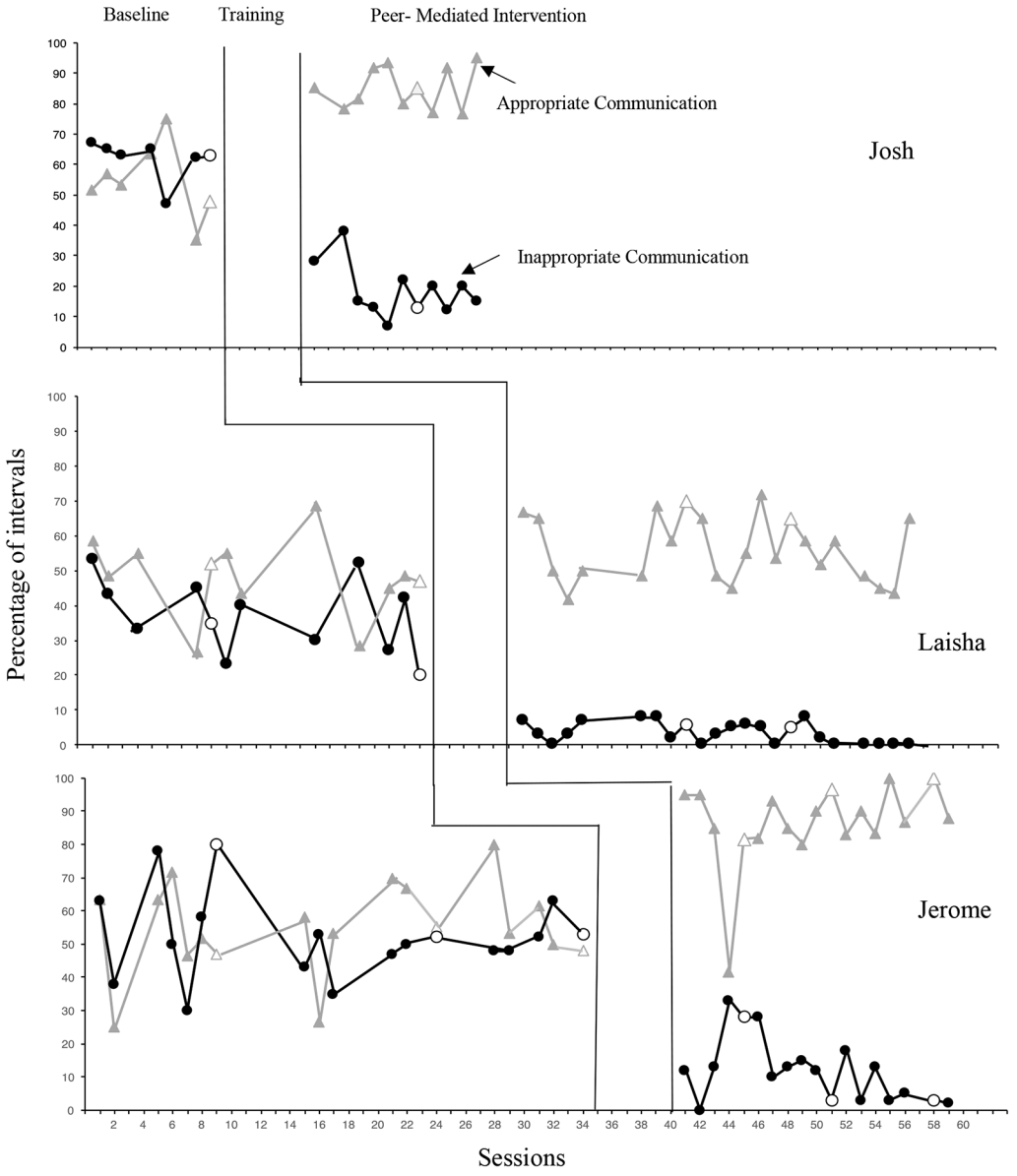


Figure 1. Percentage of intervals with appropriate (closed triangles) and inappropriate (closed circles) communication acts with trained network peers. Open symbols indicate generalization probes with untrained peers.

effect for inappropriate communicative acts (1.0) and minimal effect (with regard to non-overlap with baseline) for appropriate communicative acts (0.30).

During baseline, Jerome's inappropriate communicative acts (M = 50.5) were highly variable and higher than his appropriate acts (M = 56). At the onset of PMI, Jerome's inappropriate communication acts (M = 11), immediately decreased in level and trend, with near or at zero rates occurring during the last two sessions. Appropriate communication acts, immediately increased with intervention, and except for session 44, occurred consis-

TABLE 3

Tau-U Effect Sizes for Inappropriate and Appropriate Communication Acts

	Josh	*Laisha*	*Jerome*
Inappropriate			
Tau-U	1.0	1.0	0.99
C.I. 90%	0.49, 1.0	0.63, 1.0	0.64, 1.0
p-value	<.001	<.001	<.001
Appropriate			
Tau-U	1.0	0.30	0.88
C.I. 90%	0.49, 1.0	0.08, 0.67	0.53, 1.0
p-value	<.001	0.19	<.001

Note: C.I. = Confidence Interval.

tently above baseline rates (M = 87). Tau-*U* for Jerome's inappropriate and appropriate communication acts indicated a large (0.99) and moderate (0.88) intervention effect, respectively.

Generalization

For all focus students, generalization probe data displayed in Figure 1 (open symbols) were comparable to their session performance during both baseline and intervention conditions. For Josh and Jerome, change in the probe measures, relative to baseline, occurred during PMI indicating that decreases in inappropriate and increases in appropriate communication acts generalized to untrained peers as a result of the intervention. For Laisha, generalization improvements with untrained peers occurred only for inappropriate communications acts.

Strategy Use

Data showing overall mean percentage of intervals in which peers and focus students used the conversation strategies are presented in Table 4. The most frequently used strategies by peers were *Show Interest* and *Keep the Conversation Going*; the least frequently used strategies were *Pause/Prompt to Initiate*, *Help Your Friend Respond*, and *Redirect/Move On*. However, peers' use of the strategies varied among the focus students. For example, Josh's peers used more redirection, while Laisha's peers rarely redirected. Conversely, the focus students used their conversation cards frequently to initiate and to maintain the conversation by making statements or asking questions about the topics depicted on the card.

Social Validity

Peers. Peers rated all 11 items on the acceptability survey positively, with averages for each item exceeding 5.4 on the 7-point scale. Peers rated two items pertaining to implementing the PMI strongly positive (understanding the intervention, M = 6.6, and liking the intervention, M = 6.1), while ease of using the PMI was rated slightly less positive (M =

TABLE 4

Overall Mean Percentage of Intervals In Which Peers and Focus Students Used Conversational Strategies

	Paused and/or Prompted to Initiate		*Showed Interest*		*Kept the Conversation Going*		*Helped the FS Respond*		*Redirected/ Moved On*	
Peer Network										
Josh	23	(0-50)	100	100	100	100	29	(0-70)	56	(20-80)
Laisha	18	(0-40)	99	(80-100)	82	(50-100)	2	(0-10)	1	(0-10)
Jerome	9	(0-30)	100	100	95	(80-100)	14	(0-50)	13	(0-30)
Focus Student	Initiated using the Card		Maintained using the Card							
Josh	26	(10-50)	78	(60-90)						
Laisha	42	(20-50)	76	(60-100)						
Jerome	34	(20-50)	56	(30-100)						

Note: FS = Focus Student.

5.4). Notably, peers unanimously agreed that their training improved conversations with the focus student ($M = 7.0$). Additionally, all peers indicated that the PMI helped to improve their ($M = 6.3$) and the focus student's conversational skills ($M = 6.1$). They indicated that they would most likely recommend the PMI to a friend ($M = 6.3$) and believed the PMI resulted in positive benefits for themselves ($M = 6.3$) and the focus student ($M = 6.3$). On the other hand, the peers rated their comfort with implementing the PMI ($M = 5.7$) and things they did not like about PMI ($M = 5.7$, negatively worded) as slightly less positive. Peer write-in responses suggested that these lower ratings reflected their discomfort with the focus students' inappropriate talking during baseline.

Focus Students. With few exceptions, focus students rated all five items pertaining to their enjoyment of talking with new friends (i.e., peers), the perceived enjoyment of their friends, the similarity of conversations with other classmates, their comfort with talking with their new friends, and ease of having conversations with their friends positively. Jerome indicated he was highly satisfied, rating all items a 5. Josh and Laisha rated all items a 4 or 5, except for one item each ("I have similar conversations with my classmates" for Josh and "It is easy to have lunchtime conversations with my friends" for Laisha), which were rated a 3 ("so-so").

Educators' ratings. Educators' ratings for the quality of conversational interactions between the focus students and network peers are presented in Table 2. The educators noted substantial improvements from baseline to intervention for both focus student and peer conversational behaviors, and the overall quality of conversational interactions, indicating that intervention conversations were relaxed and natural. Post training ratings averaged above a 4.1 on the 5-point scale for all items.

Discussion

This study evaluated the effects of a novel, multicomponent PMI on the social conversation of high school students with ASD who engaged in high rates of inappropriate communication acts. The results were promising in that all focus students made substantial gains in conversing with peers in a natural high school setting. First, the findings suggest that the PMI was strongly effective in reducing the focus students' high rates of inappropriate communications, resulting in near zero levels for Laisha and Jerome, who were exposed to the intervention the longest. Appropriate communication acts consisting of appropriate initiations and topic maintenance responses increased for Josh and Jerome. Laisha's rate of appropriate communication, became more consistent with PMI, but did not increase appreciably over baseline; however, with the near elimination of inappropriate talking, the overall quality of her conversation improved.

Second, the PMI resulted in the focus students generalizing their improved conversation gains with untrained peers with the aid of their conversation card. And third, the PMI yielded socially validated outcomes. Network peers rated the intervention and conversational outcomes positively. Focus students indicated they enjoyed talking with their new friends and felt comfortable with their interactions. Additionally, educators who were naïve to the intervention noted considerable improvement in the quality of conversational interactions and growth in the focus students' conversational abilities (i.e., responsiveness, reciprocal turn-taking, and interest in peers), indicating that the PMI had a qualitative and meaningful impact.

This study contributes to the PMI literature in two important ways. First, it adds to the relatively small number of PMI studies conducted in high school focused specifically on improving the conversational interactions of adolescents with ASD with their peers. (e.g., Bambara, Cole, et al., 2016; Hughes et al., 2013; Hughes et al., 2011). This study provides additional evidence that, with minimal peer and focus student training and with the aid of visual cues, improvements in conversation can be achieved without direct adult intervention in a natural setting. This finding holds strong promise for improving social communication across the school day wherever peer-to-peer conversation is appropriate. Second, this study continues to document that PMI can be modified to address specific pragmatic targets for students with ASD (e.g., Bambara, Cole, et al.,

2016; Kamps et al., 2014; Thiemann & Goldstein, 2004). Several studies have examined the relation between improved social interaction with peers and the reduction of inappropriate social behavior (e.g., stereotypic behavior) by children with autism (e.g., Lee et al., 2007; Loftin et al., 2008). This study's unique contribution was adapting a PMI to reduce high rates of inappropriate communication acts within the context of social conversation. To date, PMI research has been primarily concerned with increasing social-communication interactions or skills, with little attention devoted to directly reducing or assessing inappropriate communication acts common among many individuals with ASD. Left unaddressed, these common conversational difficulties such as abrupt topic changes, topic perseveration, irrelevant and vague comments and poor responsiveness (Paul, 2009) can substantially interfere with the quality of conversational interactions and peer acceptance. Despite the focus students' high rates of inappropriate talking during baseline, peers reported after the intervention that they enjoyed spending time with the focus students would recommend the intervention to a friend. This positive outcome highlights the importance of finding effective interventions for students with ASD who present difficult communication challenges.

The PMI was introduced as a package, making it impossible to determine which component, the focus students' use of their conversation card or the peers use of their conversational support strategies, contributed most to the outcomes. However, attempting to separate out the relative contributions of the peer and focus student components in PMI may be undesirable because each plays a vital role in influencing conversational interactions. The data show that the focus students made frequent use of their conversation cards which, consistent with research investigating the use of text or graphic cues for conversation (e.g., Koegel et al., 2014), provided a visual structure and efficient means for engaging in appropriate on-topic conversations with peers. Trained to be responsive partners, the data also show that network peers encouraged appropriate conversation by showing interest in the focus students' communications, extending interactions by asking questions or sharing information, prompting the use of the conversation card and redirecting inappropriate or off-topic communication acts as needed. Recent research investigating the role of peers in PMI packages suggests that training peers as responsive partners positively influences the rate with which students with ASD will use newly trained communication skills and visual aids in natural social contexts (Schmidt & Stichter, 2012; Thiemann-Bourque et al., 2016). Differences in the peer networks' use of the strategies, especially prompts and redirection, may be partially explained by the degree of support needed by the focus students to use their cards and engage appropriate conversations. We believe that generalization with untrained peers occurred not only because the conversation card helped to mediate generalization, but also because the responses of the trained network peers brought the focus students' appropriate conversation under natural contingencies of reinforcement.

Focus students responded to the intervention rather quickly. There may be several reasons for this. First, as evidenced by their high rates of talking, focus students were motivated speakers. Consistent with how verbal noncommunicators have been described in the language literature (e.g., Fey, 1986; Fujiki & Brinton, 1991), they appeared to use both appropriate and inappropriate acts to maintain the speaker floor. Similarly, several researchers (Fisher et al., 2013; Rehfeldt & Chambers, 2003) have documented that the perseverative speech of some individuals with ASD functions to secure partner attention. With this understanding, use of the conversation card, reinforced and supported by peers, may have functioned as an effective replacement for inappropriate talk to secure peer attention. Fisher et al. (2013) similarly reported that pairing visual stimuli to signal appropriate topics combined with differential reinforcement effectively reduced the perseverative speech and increased appropriate turn-taking by an adolescent with ASD in a clinical setting. Second, constructing the conversation card around the focus students' interests may have provided additional motivation for engaging in appropriate conversation. Multiple studies have shown that improved social interaction can be achieved by incorpo-

rating the preferred interests of individuals with ASD in the intervention (Gunn & Delafield-Butt, 2016).

The results of this novel use of PMI with students who present difficult conversational challenges is encouraging, however, more research is needed to address the study's limitations and build upon its findings. First, the PMI impacted the focus students' rate of appropriate communication acts differentially. One potential reason for students' differential response to the intervention, may be explained by the differences in their communication styles. We observed that during intervention, Laisha tended to ask each peer a question when initiating a topic, giving each an opportunity to respond before she shared information about the topic herself. This resulted in longer conversational exchanges with peers and less overall talking by Laisha once inappropriate acts decreased. Josh and Jerome on the other hand, simply talked more, albeit appropriately about the topics on their cards during intervention. In addition to evaluating overall rates of appropriate and inappropriate talking, future research would benefit from measuring the specific communication acts of both the focus students and peers to more precisely document how the intervention impacts various communicative behaviors (e.g., initiations, commenting, follow-up questions) and reciprocal turn-taking (e.g., the number of turns within a topic). Such an analysis would provide more precise information on how each participant responded to the intervention, such as in Laisha's case, and lead to refining future iterations of the intervention.

Second, although topics on the conversation cards varied across sessions and successfully prevented topic repetition or perseveration within a session, some peers found many of the focus students' limited interests (e.g., certain videogames, tv shows) boring over time. Future researchers should consider ways of teaching students with ASD how to balance conversational topics and turn-taking between their own and their peers' interests (e.g., Fisher et al., 2013). Lastly, although we assessed generalization with untrained peers, additional research is needed to evaluate the full extent of generalization (e.g., across different settings, without the use of the conversation card) and to extend improved interactions across naturally occurring social opportunities throughout the day.

In summary, the conversational difficulties of high school students with ASD can limit peer interactions and social relationships. The results of this initial investigation document that peer mediated intervention can be adapted to effectively improve the social conversation of adolescents characterized as nonverbal communicators during lunch with trained and untrained peers. Inappropriate communication acts were substantially reduced, and importantly, the intervention produced socially validated outcomes, attesting to the quality of the conversations and the feasibility of the intervention. This study provides promising evidence for the utility of the intervention in high school settings and for future research.

References

Bambara, L. M., Chovanes, J., Thomas, A., & Cole, C. L. (2016). Effective peer-mediated strategies for improving the conversational skills of adolescents with autism. *Perspectives of the ASHA Special Interest Groups, 1,* 29–36.

Bambara, L. M., Cole, C. L., Chovanes, J., Telesford, A., Thomas, A., Tsai, S. C., Bilgili, I. (2018). Improving the assertive conversational skills of adolescents with autism spectrum disorder in a natural context. *Research in Autism Spectrum Disorders, 48,* 1–16.

Bambara, L. M., Cole, C. L., Kunsch, C., Tsai, S. C., & Ayad, E. (2016). A peer-mediated intervention to improve the conversational skills of high school students with autism spectrum disorder. *Research in Autism Spectrum Disorders, 27,* 29–43.

Carter, E. W., Hughes, C., Guth, C. B., & Copeland, S. R. (2005). Factors influencing social interaction among high school students with intellectual disabilities and their general education peers. *American Journal on Mental Retardation, 110,* 366–377.

Chung, K. M., Reavis, S., Mosconi, M., Drewry, J., Matthews, T., & Tassé, M. J. (2007). Peer-mediated social skills training program for young children with high-functioning autism. *Research in Developmental Disabilities, 28,* 423–436.

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

Fey, M. E. (1986). *Language intervention with young children.* College-Hill Press.

Fisher, W. W., Rodriguez, N. M., & Owen, T. M. (2013). Functional assessment and treatment of perseverative speech about restricted topics in an adolescent with Asperger syndrome. *Journal of Applied Behavior Analysis, 46,* 307–311.

Fujiki, M., & Brinton, B. (1991). The verbal noncommunicator: A case study. *Language, Speech, and Hearing Services in Schools, 22,* 322–333.

Goldstein, H., Schneider, N., & Thiemann, K. (2007). Peer-mediated social communication intervention: When clinical expertise informs treatment development and evaluation. *Topics in Language Disorders, 27,* 182–199.

Gunn, K. C., & Delafield-Butt, J. T. (2016). Teaching children with Autism spectrum disorder with restricted interests: A review of evidence for best practice. *Review of Educational Research, 86,* 408–430.

Harrison, J. R., State, T.M., Evans, S. W., & Schamberg, T. (2016). Construct and predictive validity of social acceptability scores from high school teacher ratings on the School Intervention Rating Form. *Journal of Positive Behavior Interventions, 18,* 111–123.

Hughes, C., Bernstein, R. T., Kaplan, L. M., Reilly, C. M., Brigham, N. L., Cosgriff, J. C., & Boykin, M. P. (2013). Increasing conversational interactions between verbal high school students with autism and their peers without disabilities. *Focus on Autism and Other Developmental Disabilities, 28,* 241–254.

Hughes, C., Golas, M., Cosgriff, J., Brigham, N., Edwards, C., & Cashen, K. (2011). Effects of a social skills intervention among high school students with intellectual disabilities and autism and their general education peers. *Research and Practice for Persons with Severe Disabilities, 36,* 46–61.

Hughes, C., Rung, L. L., Wehmeyer, M. L., Agran, M., Copeland, S. R., & Hwang, B. (2000). Self-prompted communication book use to increase social interaction among high school students. *Research and Practice for Persons with Severe Disabilities, 25,* 153–166.

Kamps, D., Mason, R., Thiemann-Bourque, K., Feldmiller, S., Turcotte, A., & Miller, T. (2014). The use of peer networks to increase communicative acts of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 29,* 230–245.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Koegel, L. K., Park, M. N., Koegel, R. K (2014). Using self-management to improve the reciprocal social conversation of children with autism spectrum disorder. *Journal of Autism and Developmental Disabilities, 44,* 1055–1063.

Kucharczyk, S., Reutebuch, C. K., Carter, E. W., Hedges, S., El Zein, F., Fan, H., & Gustafson, J. R. (2015). Addressing the needs of adolescents with autism spectrum disorder: considerations and complexities for high school interventions. *Exceptional Children, 81,* 329–349.

Lee, S., Odom, S. L., & Loftin, R. (2007). Social engagement with peers and stereotypic behavior of children with autism. *Journal of Positive Behavior Interventions, 9,* 67–79.

Locke, J., Ishijima, E. H., Kasari, C., & London, N. (2010). Loneliness, friendship quality and the social networks of adolescents with high-functioning autism in an inclusive school setting. *Journal of Research in Special Educational Needs, 10,* 74–81.

Loftin, R. L., Odom, S. L., & Lantz, J. F. (2008). Social interaction and repetitive motor behaviors. *Journal of Autism and Developmental Disorders, 38,* 1124–1135.

Paul, R. (2008). Interventions to improve communication in autism. *Child and Adolescent Psychiatric Clinics of North America, 17,* 835–856.

Paul, R., Orlovski, S. M., Marcinko, H. C., & Volkmar, F. (2009). Conversational behaviors in youth with high-functioning ASD and Asperger syndrome. *Journal of Autism and Developmental Disorders, 39,* 115–125.

Rehfeldt, R. A., & Chambers, M. R. (2003). Functional analysis and treatment of verbal perseverations displayed by an adult with autism. *Journal of Applied Behavior Analysis, 36,* 259–261.

Schmidt, C., & Stichter, J. P. (2012). The use of peer-mediated interventions to promote the generalization of social competence for adolescents with high-functioning autism and Asperger's Syndrome. *Exceptionality, 20,* 94–113.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale – 2nd Edition* (CARS2). Western Psychological Services.

Thiemann, K. S., & Goldstein, H. (2001). Social stories, written text cues, and video feedback: Effects on social communication of children with autism. *Journal of Applied Behavior Analysis, 34,* 425–446.

Thiemann, K. S., & Goldstein, H. (2004). Effects of peer training and written text cueing on social communication of school-age children with pervasive developmental disorder. *Journal of Speech, Language, and Hearing Research, 47,* 126–144.

Thiemann-Bourque, K., Brady, N., McGuff, S., Stump, K., & Naylor, A. (2016). Picture Exchange Communication System and Pals: A peer-mediated augmentative and alternative communication intervention for minimally verbal preschoolers with autism. *Journal of Speech, Language, and Hearing Research, 59,* 1133–1145.

Vannest, K.J., Parker, R.I., Gonen, O., & Adiguzel, T. (2016). Single case research: Web based calculators for SCR analysis. (Version 2.0) [Web-based application]. Texas A&M University. Retrieved August 16, 2018. Available from singlecaseresearch.org

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., .. Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders, 45,* 1951–1966.

Received: 19 December 2018
Initial Acceptance: 13 February 2019
Final Acceptance: 8 June 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 201–214

Effectiveness of the Parent Training Program for Supporting the Preparation of Individuals with Intellectual Disability for Adulthood on Mothers' Quality of Life Perceptions

Gizem Yildiz and Atilla Cavkaytar
Anadolu University

Abstract: The education offered for the parents of individuals with intellectual disability improves the quality of life of the families, and thus the quality of life of the individuals with intellectual disability in the long term. The aim of the present study was to investigate the effects of the Parent Training Program for Supporting the Preparation of Individuals with Intellectual Disability for Adulthood (PSIDA) that was developed based on the requirements of the families of individuals with intellectual disability on the quality of life perceptions of the families. The families of adults with intellectual disability require support concerning adaptation to adulthood, preparation for the future, coping with future stress, social support, employment and legal rights. PSIDA was developed based on these requirements. The study was conducted with a pre-test/post-test control group design. Eighteen mothers (nine experimental, nine control group) participated in the study. The findings demonstrated that there was a significant difference between the experimental and control groups, and the quality of life perceptions of the parents that participated in the training improved. This improvement was observed in the subdomains of parenting and emotional well-being dimensions in the Beach Center Family Quality of Life Scale.

Family Quality of Life of Adults with Intellectual Disability

Individual differences make each person unique. These differences can be observed in personal, intellectual, social, emotional or physical developmental areas. Individuals with special needs differ significantly from their peers based on individual and developmental characteristics (Friend, 2011). Those who differ in the domain of intellectual development are defined as individuals with intellectual disability (ID). "*Intellectual disability* is a disability characterized by significant limitations in both intellectual functioning and in adaptive behavior, which covers many everyday social and practical skills. This disability emerges before the age of 18" (AAIDD, 2019). New approaches emphasized that individuals with ID should be evaluated in multidimensionality, not only in certain dimensions, and accordingly, different support requirements should be determined. An important aspect of these support requirements is preparation for adulthood. The components of the Support Intensity Scale-Adult Version (SIS-A) determine the requirements needed during the stage of preparation for adulthood. Previous studies demonstrated that the requirements of individuals with ID, who prepare for adulthood, were related to daily life, community life, lifelong learning, employment, health and safety, and social adaptation (Thompson et al., 2004; Thompson et al., 2015). Thus, it can be suggested that the requirement-based support provided to adults with ID would increase their quality of life, and thus that of their families (Schalock & Verdugo, 2002).

The quality of life of individuals with ID and their full participation in society and independence are directly associated with their preparation for the post-schooling training period, professional life and adulthood (Alwell & Cobb, 2009; Drew & Hardman, 2007; Turnbull et al., 2013; Wehman, 2013). While individuals with ID receive training to develop

Correspondence concerning this manuscript should be addressed to Gizem Yildiz, Anadolu University, Faculty of Education, Department of Special Education, Eskisehir, 26210, TURKEY. E-mail: gizemy@anadolu.edu.tr and 88.gizemyildiz@gmail.com

independent living skills during formal education, they start to experience difficulties in their participation in adult and social life after the school period (Will, 1983). Previous studies demonstrated that several parents do not know what will happen after school is over and worry about the future (Davies & Beamish, 2009; Griffin et al., 2010; Kraemer & Blacher, 2001; Rossetti et al., 2016).

These difficulties experienced by individuals with ID during transition to independent living and adult life undoubtedly affect family members in different ways (Davies & Beamish, 2009; Grant & Whittell, 2001; Mclntyre et al., 2002). The negative experiences of children with ID in this period could reduce the quality of life perceptions of their parents (Chou et al., 2007; Walden et al., 2000; Yoong & Koritsas, 2012). The quality of life is defined as the subjective living conditions of the individual and the subjective self-perception of the individual related to these conditions such as fulfillment of personal needs in social groups such as family, school and work, and fulfillment of basic responsibilities (Keith et al., 1996; Schalock, 1994). The concept of quality of life of the family was derived from the quality of life, manifested in various social elements such as psychology, health, politics, and disability (Bayat, 2005). Family Quality of Life is the ability of family members to meet their needs, to do the things that are important to them, and to enjoy living together as a family (Meral & Cavkaytar, 2013; Park et al., 2003; Turnbull et al., 2000).

Participation of an individual with ID in independent life is a multi-dimensional process. The family has the most significant role among these dimensions. So much so, that the cultural, material, social, emotional and cognitive familial conditions have an impact on the life of the individual with ID (Smith et al., 2006). Thus, it can be suggested that improving the perception of the quality of life of the family leads indirectly to an increase in the perception of quality of life of the individual.

The quality of life of families with children with ID could be affected by the level of disability, experiences and difficulties of the child. As mentioned above, families need support especially during the preparation for adulthood after the school years (Davies & Beamish, 2009; Griffin et al., 2010; Rossetti et al., 2016). Youth and adulthood years, which span the post-school period, could be identified as a new era in which independent life begins (Arnett, 2010; Wehman, 2013). The problems experienced by families in this period that includes the transition to adulthood concentrate on topics such as future anxiety, the expectation for school attendance, lack of knowledge of adult services, conflicts between the desires of the parents and children, employment and retention, vocational training requirements, and marriage expectations (Blacher, 2001; Steere et al., 2007; Stewart, 2009; Turnbull et al., 2006).

A literature review demonstrated that the vast majority of studies on family quality of life of individuals with ID were conducted to determine the family quality of life perceptions and their views on those perceptions (Brown et al., 2006; Chou et al., 2007; Hoffman et al., 2006; Jokinen & Brown, 2005; Mactavish et al., 2007; Mugno et al., 2007; Park et al., 2003; Poston et al., 2003; Vanderkerke et al., 2018; Viana Tomaz et al., 2017; Walden et al., 2000; Werner et al., 2009). On the other hand, a few studies included implementations to inform and support parents of individuals with ID on topics such as preparing for adulthood and the future, reducing stress, improving emotional and self-efficacies, improving perceptions of social support, and parenting and family interaction supports (Cankaya, 2013; Cavkaytar, 1999; Cavkaytar et al., 2014; Coren et al., 2018; Fodstad et al., 2018; Hand et al., 2012; Hudson et al., 2003; Martinez-Rueda et al., 2017; Yildiz & Cavkaytar, 2017). Only two studies aimed to investigate the impact of family training programs on the quality of life perception of families (Cavkaytar et al., 2014; Martinez-Rueda et al., 2017).

In a single-subject study conducted by Cavkaytar (1999) with mothers of children with ID, a parent training program was offered, and it was determined that mothers could help their children to acquire self-care and daily lifeskills. In an empirical study conducted by Smith et al. (1996) on the effectiveness of a parent training program provided to 65 years and older parents and developed to provide psychological support and knowledge on coping with age-related difficulties, recognizing the legal rights of their children, supporting each other and improving solidarity, it

was determined that the program was effective and all parents reported positive views on the program.

Solomon et al. (1997) investigated the effects of two family training programs for individuals with ID on the stress and self-efficacy of the families. Although the study findings reflected no significant difference between the experimental and control groups, it was reported that the three-month training period improved self-efficacy and self-esteem of the groups that received the training. Hudson et al. (2003) conducted a quasi-experimental study where three modes of flexible application systems (group support, telephone support and a series of instruction manuals) were used based on the guides they developed for families. The study findings demonstrated that the stress levels of parents who participated in three applications were lower when compared to those of families who did not participate. The first group exhibited a higher performance in parenting roles and there was no significant difference between the groups that attended the three training programs.

In an empirical study by Hand et al. (2012) that investigated the impact of a family training program on the psychological distress and difficulty levels of the families, it was determined that the program was effective on improving the strengths of the families and their self-efficacy levels and reducing stress. In another study, a mobile skills instruction program for family members to help adults with ID to acquire daily life skills was developed and its effectiveness was tested. In that single-subject study, as a result of the instructional activities carried out by the family members using the software, adults with ID were able to learn daily life skills (Cankaya, 2013). Yildiz and Cavkaytar (2017), in their empirical study, discovered that the sexual education program developed for the mothers of adults with ID positively changed the attitudes of mothers towards sexual education and improved their perceptions about social support. In a literature review conducted by Coren et al. (2018), family training implementations for the parents of individuals with ID were evaluated based on parent support, improvement of parent-child interaction, preservation of the family environment and development of parenting skills variables. Thus, in empirical studies on family training conducted until 2017, it was determined that the participants were mostly mothers, the majority of the studies were for parents with young children, the family training reduced the parental stress and developed parenting skills, while a few studies reported on the development of the family training programs, the program content and who applied the program.

A review of empirical studies that investigated the quality of life perception of families who have children with ID revealed that Cavkaytar et al. (2014) determined the needs of parents of children with ID and analyzed the effects of the online Family Information and Support Training Program (E-FISTP) that they developed on family knowledge level and quality of life perceptions. In that empirical study, it was observed that family training increased the knowledge level of the participants, perceptions about social support and self-efficacy perceptions, however, it did not have any effect on the quality of life perceptions. In the study, it was suggested to plan effectiveness studies with programs that aim to improve the quality of life perceptions of parents in future research. Martinez-Rueda et al. (2017) developed a program to support the quality of life perceptions of families of adults with ID. In that study, where the family support program they developed was piloted with five parents, it was reported that the program was effective on increasing the quality of life of the parents. Furthermore, it was suggested to investigate the effect of family training on families' quality of life perceptions in future studies.

As detailed above, the number of empirical studies that aimed to improve the quality of life perceptions of parents of individuals with ID in the literature is very limited (Cavkaytar et al., 2014; Martinez-Rueda et al., 2017). It can be argued that the present study would fill this gap as recommended by previous studies in the literature. Previous studies demonstrated that family training was effective on many areas such as reducing parental stress, increasing quality of life, increasing self-efficacy and social support perception, increasing parenting skills, teaching skills to children, and strengthening psychological and social aspects (Cankaya, 2013; Cavkaytar, 1999; Cavkaytar et al., 2014; Coren et al., 2018; Fodstad

et al., 2018; Hand et al., 2012; Hudson et al., 2003; Martinez-Rueda et al., 2017; Smith et al., 1996; Solomon et al., 1997; Yildiz & Cavkaytar, 2017). On the other hand, it was emphasized that it is important to support families in order to improve their quality of life, to reduce their future anxiety and support should be provided to keep their desires and expectations and their future expectations realistic, information should be provided about their legal rights in adulthood, and how they can reach these services (Feldman et al., 1999; Gauthier-Boudreault et al., 2017; Knowles et al., 2015; Walden et al., 2000). Furthermore, the independence of adults with ID is in interaction with the family quality of life (Yoong & Koritas, 2012). Thus, all education that aims to improve the family quality of life would also make individuals more independent.

Research on families' quality of life perceptions are generally descriptive studies. Only two studies have examined the impact of family training programs on the families' quality of life perceptions (Cavkaytar et al., 2014; Martinez-Rueda et al., 2017). Furthermore, the determination of the needs of families when developing family support training programs is another issue (Cavkaytar et al., 2014; Gauthier-Boudreault et al., 2017; Kim & Turnbull, 2004; White & Hastings, 2004). Providing training for parents based on their needs and inclusion of parents in support of adults with ID to conduct independent lives is the duty of specialists (Sileo & Prater, 2011).

In this respect, a study that develops a family training program based on the needs of families to prepare their children with ID for adulthood and which might affect the perception of family quality of life has not been found. The aim of the present study was to investigate the effects of the Parent Training Program for Supporting the Preparation of Individuals with Intellectual Disability for Adulthood (PSIDA) based on the needs of parents of adults with ID on their quality of life perceptions.

Method

Research Design

The study was conducted with pre-test post-test control group design, an experimental research method. The Beach Center Family Quality of Life Scale was applied to the participants as pre-test and post-test. The independent variable of the study was the PSIDA, which was developed by the authors. The dependent variable was the family quality of life perceptions.

Development of PSIDA

PSIDA was developed based on the parent training program development stages as follows:

1. *Determination of parents' needs.* The first step in the development of the parent training program is the determination of needs. Family Needs Survey (FNS) was applied to all participants (n = 18) to determine the family needs. FNS was completed with face-to-face interviews with all participants in the experimental and control groups.

At this stage, the data obtained with the FNS form were analyzed with descriptive analysis and the needs were ranked based on priority. Thus, the most frequently marked items by the participants and the responses to the open-ended questions were considered as needs. These needs are presented in Table 1.

As seen in Table 1, it was determined that families required information on parenting individuals with ID, the future of their children, adult life, social life preparation, future institutions, legal rights and employment, as well as daily life and non-familial social support and psychological support.

Based on the needs presented in Table 1, the authors decided which topics would be included in the family training program. Thus, the PSIDA that included four modules [(I) The Process of Adaptation to Adulthood, (II) Social Support, (III) Coping with Stress, and (IV) Vocational Training and Employment] was developed.

Support Intensity Scale-Adult Version (SIS-A), which was translated to Turkish language and culture, though not tested for validity and reliability, is used to determine the support requirements of adults with ID in Turkey. In studies where SIS-A was implemented, it was observed that the needs of adults with ID were related to daily life, social life, lifelong learning, employment, health and safety and social

TABLE 1

The Needs of the Parents of Adults with ID During the Process of Preparation for Adulthood

The Most Frequently Marked Items in FNS	*Frequency (f)*
Item 16: I need reading material about other parents who have a child similar to mine.	17
Item 23: Our family needs help in learning how to support each other during difficult times.	15
Item 21: I need more information about the services that my child might receive in the future.	15
Item 26: I need to have someone in my family that I can talk to more about problems.	13
Item 12: I need help in knowing how to respond when friends, neighbors, or strangers ask questions about my child's condition.	12
Item 17: I need more information about my child's condition or disability.	12

The Themes Obtained with the Open-Ended Questions in FNS
Adult life and independent life
Social support
Legal Rights in adulthood
Preparation for the future and future anxiety and stress
Vocational training and employment

adaptation (Thompson et al., 2004; Thompson et al., 2015). In the present study, it can be argued that the requirements determined with FNS applied to the parents of adults with ID were consistent with the requirements determined with SIS-A for adults with ID.

2. *Introduction of the program.* PSIDA is a parent training program for institution-based groups that included four 1-hour long sessions and supported by audiovisual and written material, developed by the author to support the parents of adults with ID in adaptation to the process of transformation from post-educational life to adulthood, to assist them in creating daily life and non-familial support, to inform them about their legal rights, to assist them to cope with the process of transformation to adulthood and inform them on vocational training and employment.
3. *The aim of the program.* The aim of PSIDA is to assist the parents of adults with ID to adapt to the process of transition to adulthood, and to inform the parents on social support, vocational training, employment, legal rights and services in adulthood.
4. *The content of the program.* PSIDA consists of four sessions, namely (I) Adaptation to Adulthood, (II) Social Support, (III) Coping with Stress and (IV) Vocational Training and Employment. The program content was based on the data obtained with FNS applied to all participants.
5. *Teaching-learning processes.* The presentation of the content was conducted with PowerPoint presentation supported by audiovisual elements, cartoons and videos and the training was presented by the first author with instruction, question-answer, demonstration and implementation and animation techniques. Furthermore, a family handbook that included separate booklets for each module was developed and distributed to parents in the beginning of the session. While the sessions were conducted in the form of group training, individual interviews were conducted on the topics when needed.
6. *Testing.* In order to analyze the impact of the program on family quality of life perceptions, Beach Center Quality of Life Scale was applied to all participants before and after the implementation. Furthermore, after the training, participation certificates were awarded to the experimental group participants regarding their participation in the training.

Participants

The study participants included 18 mothers whose children with ID were employed in an assembly workshop in Eskisehir province, Turkey. Mothers were assigned randomly to experimental (nine mothers) and control (nine mothers) groups by random program (Research Randomizer, 2016).

The mothers were between 42 and 68 years old and the children with ID were between 20 and 35 years old, and all mothers were housewives. Analysis of the education levels of the participants demonstrated that 88% of mothers were primary school graduates and 22% were middle school graduates. Furthermore, all families except one had social security benefits, the average monthly income of the households was between ₺ 850 and 2000. Thus, the participants had low education and socio-economic levels.

Setting

This study was conducted within the scope of the Assembly Work Vocational Education Program (AWVEP) project carried out in Eskisehir Tepebasi Municipality İbrahim Ethem Kesikbas Assembly Workshop for Individuals with Disabilities with the support of Turkish Employment Organization (ISKUR) and Anadolu University. In this context, the parent training was carried out in the meeting hall at the upper floor of the Eskisehir Tepebasi İbrahim Kesikbas Assembly Workshop where the adults with ID were employed. The hall is about 15 square meters and there is a meeting table and chairs for 15 individuals in the hall. In parent training sessions, the mothers were seated to form a U and the author conducted the presentation with a projection device. Parent training sessions were conducted by the first author. All sessions were recorded.

Data Collection

Three quantitative data collection techniques were employed in the study. Initially, a questionnaire was utilized during the determination of the needs, a quality of life scale was used to test the effectiveness of PSIDA, and finally, a checklist was used to assess the reliability of the implementation.

The Family Needs Survey (FNS) was applied to all participants in order to determine the needs of the families. FNS is a 3-point Likert type questionnaire that includes 29 items and it was developed by Bailey and Simeonsson (1988) and adapted to Turkish by Bulbin Sucuoglu (1995). The data collected at this stage were used to develop the PSIDA content.

The study effectiveness data were collected with Beach Center Family Quality of Life (BCFQOL) scale. The scale, developed by University of Kansas Beach Center on Family and Disability (2006) and adapted to Turkish by Meral (2012), is a data collection instrument that includes 25 questions, five subdomains and 5 point likert type answers and aims to determine the quality of life of families with children with special needs (Meral & Cavkaytar, 2013). The subdomains of BCFQOL include family interaction, parenting, emotional well-being, physical/material well-being, and disability-related support. The Beach Center Family Quality of Life Scale used in the study is a 5 point Likert-type scale and it is scored as follows; completely not suitable: 1 point, not suitable: 2 points, neither suitable nor unsuitable: 3 points, suitable: 4 points, completely suitable: 5 points. Thus, the minimum possible score in the scale is 25 points and the maximum score is 125 points. The score obtained by dividing the total scale score obtained by the number of items allows for comment on the quality of life perception levels. Accordingly, for this 5-point Likert type scale; 1 is the lowest, 2 is a low, 3 is a moderate, 4 is a high and 5 is the highest score and the quality of life perceptions of those who score 3 points and below are considered as low and of those who score 3 and above are considered as high.

The implementation reliability data for the program was collected with "PSIDA Implementation Evaluation Checklist" developed by the first author and independent observer in the field of special education. The PSIDA Evaluation Checklist is a 16-item checklist developed by the author to check whether the implementation process was conducted as planned.

Procedure

Parent training sessions. PSIDA is a parent training program for institution-based groups that included four 1-hour long sessions and

supported by audiovisual and written material, developed by the author to support the parents of adults with ID in adaptation to the process of transformation from post-educational life to adulthood, to assist them in creating daily life and non-familial support, to inform them about their legal rights, to assist them to cope with the process of transformation to adulthood and inform them on vocational training and employment.

PSIDA was implemented as a group training that included four 1-hour long sessions. All group training sessions lasted one week in December, 2016. At the beginning of the first session and at the end of the last session, participants took the quality of life scale as pre-test and post-test. The Process of Adaptation to Adulthood topic was instructed in the first session, the Social Support topic was instructed in the second session, the Coping with Stress topic was instructed in the third session, and the Vocational Training and Employment topic was instructed in the fourth session. All sessions were instructed by the first author. Instruction, question-answer technique, and animation technique were used in the sessions and the presentations were enriched with audiovisual elements. In the final 15 minutes of each session, a discussion was conducted, and the questions, comments and suggestions of the mothers on the topics were discussed.

During the sessions, not only the content knowledge was presented, but also instructions and animations on how to plan activities to improve the quality of life, how they can organize their lives, how they can cope with stress, what they will encounter in the process of adaptation to adulthood, and how they can take precautions. Furthermore, the presentations were supported by cartoons developed in the study. A family handbook that included program content was developed and the related booklet was distributed to mothers at the beginning of each session. Activities aimed to improve the quality of life were enacted with the adults with ID who were employed in the workshop in addition to the parent training sessions to support the children of the participants. The children were asked to implement the activities and information they learned in the workshop with their families at home. This program aimed to include the children in the parent training and the children acquired knowledge consistent with their mothers. At the end of the training, a certificate of participation was awarded to the mothers who participated in the parent training sessions.

Ethics

Written consent was obtained from both the participants and Eskisehir Tepebasi Municipality where the study was conducted. Furthermore, a consent form that stipulated that the participants can withdraw from the study at any time was signed by all participants. The same training sessions was also provided for the control group after the study was completed.

Data Analysis

Descriptive analysis was conducted on the data collected with FNS and the most frequently identified items were determined as needs. Furthermore, themes were determined with qualitative descriptive analysis conducted on the open-ended question responses and included based on priority ranking. The study effectiveness data were analyzed statistically with SPSS software. Nonparametric test analysis was conducted due to the inability to reach the sufficient number of experimental and control group subjects to conduct parametric tests and due to the lack of normal distribution in all datasets. The Mann Whitney U Test was used to determine whether PSIDA led to a difference in the family quality of life perceptions. In order to conduct this test, the gain scores between the experimental and control group pre-test and post-test scores were calculated and the dependent variable was reduced to a single level. Similarly, Mann Whitney *U* Test was conducted separately for each subdomain in the quality of life scale.

Results

Effectiveness of PSIDA

To obtain the findings on the effect of the PSIDA on the participants' quality of life, pretest post-test descriptive statistics were calculated for each group and Mann Whitney *U*

TABLE 2

Descriptive Statistics on the Quality of Life Perceptions of the Parents with Children with ID

Group	*Mean*	*Std. Deviation*	*Std. Error*	*N*
Experimental				9
Pre-test score	40,00	9,01	3,00	
Post-test score	69,88	14,31	4,77	
Gained score	29,88	9,61	3,20	
Control				9
Pre-test score	43,88	10,24	3,41	
Post-test score	38,88	8,57	2,85	
Gained score	−5,00	11,39	3,80	
Total				18

Test results were calculated based on the inter-group score differences.

As seen in Table 2, the experimental group mean pre-test score was 37.22, and the post-test mean score was 81.11. The control group mean pre-test score was 43.38 and the mean post-test score was 38.88. Thus, it was observed that both the experimental group scores increased from the pre-test to post-test and the mean of experimental group post-test score higher than the control group.

When the total scale scores were divided by the number of questions, the obtained figure reflects the level of quality of life perceptions. Thus, the quality of life perceptions in the experimental group were very low (40/25 = 1.6) before the training and the quality of life perceptions improved after the training (69.88/25 = 2.80), while the quality of life perceptions in the control group were very low based on both pre-test (43.88/25 = 1.76) and post-test scores (38.88/25 = 1.56).

Since all calculations were lower than 3, the general family quality of life perceptions of all participants were lower than average based on the scale scoring principles, while there was an increase in the family quality of life perceptions of the experimental group. The results of the Mann-Whitney *U* Test that was conducted to determine whether the increase was significant are presented in Table 3.

Based on the Table 3, the difference between the experimental and control group quality of life scale scores was significant ($p < .001$). In other words, there was a significant difference between the scores of the mothers who attended and did not attend PSIDA and the quality of life perceptions of the mothers who attended the program improved as demonstrated by descriptive findings. Thus, it can be suggested that the parent training program was effective in increasing the quality of life perceptions of mothers. The results of the Mann-Whitney *U* Test conducted for each subdomain of the scale to determine the dimensions and subdomains where the above-mentioned improvement was observed are presented in Table 4.

As presented in Table 4, there was a significant difference between the quality of life scale scores of the experimental and control groups only in the subdomains of parenting

TABLE 3

Mann-Whitney U-Test Results Conducted to Compare BCFQOL Score Differences of the Parents

Group	*N*	*Mean Rank*	*Sum of Rank*	*U*	*z*	*p*
Experimental	9	14,00	126,00	.000	−3,59	.000
Control	9	5,00	45,00			
Total	18					

TABLE 4

Mann-Whitney U-Test Results Conducted to Compare BCFQOL Score Differences of the Parents Based on the Groups

Subdomain	*Group*	*Mean Rank*	*Sum of Rank*	*U*	*z*	*p*
Family interaction	Experimental	10,06	90,50	35,500	−.445	.656
	Control	8,94	80,50			
Parenting*	Experimental	14,00	126,00	.000	−3,60	.000*
	Control	5,00	45,00			
Emotional well-being*	Experimental	14,00	126,00	.000	−3,59	.000*
	Control	5,00	45,00			
Physical/material well-being	Experimental	9,06	81,50	36,500	−.359	.719
	Control	9,94	89,50			
Disability-related support	Experimental	10,61	95,50	30,500	−.891	.373
	Control	8,39	75,50			

and emotional well-being ($p < .001$). Thus, it can be suggested that the perceptions of the participants in the experimental group only increased in the subdomains of parenting and emotional well-being after the training. This finding demonstrated that PSIDA was effective in improving the family quality of life perceptions by improving the participants' perceptions in the subdomains of parenting and emotional well-being. The perception levels of the parents in each subdomain were as follows: family interaction experimental group pre-test 1.83, post-test 2.08; control group pre-test 1.96, post-test 1.85; parenting experimental group pre-test 1.62, post-test 3.62, control group pre-test 1.77, post-test 1.66; emotional well-being experimental group pre-test 1.45, post-test 3.44, control group pre-test 1.58, post-test 1.50; physical well-being experimental group pre-test 1.44, post-test 1.62, control group pre-test 1.55, post-test 1.44; and finally, disability support experimental group pre-test 1.69, post-test 2.47, control group pre-test 1.80, post-test 1.64. In this way, the subdomain perceptions that increased from lower level to the moderate level after the training were observed in parenting (from 1.62 to 3.62) and emotional well-being (from 1.45 to 3.44) subdomains.

Treatment Fidelity

The behavior of the practitioner was observed in all sessions by an independent observer and the latter assessed whether the instructor offered the program as planned. Independent variable reliability data for PSIDA were analyzed with descriptive analysis. The instructor (first author) showed 97% compliance with the steps of each experimental session.

Discussion

Based on the results of the present study, where the effects of the parent training program developed based on the needs of the parents with adult children with ID on their quality of life perceptions were scrutinized, it was determined that the parents required information and assistance in topics such as transition to adulthood, adaptation to adulthood, obtaining social support, preparing for the future, and coping with the related stress. The parent training program (PSIDA) that was developed based on the above-mentioned needs was effective in increasing the quality of life perceptions of the mothers (see Tables 2 and 3).

In the study, it was determined that parents required information on preparing for adulthood, the institutions the children would attend in adulthood, their legal rights and employment, and social support. The review of the reports on the needs of parents with children with ID in the literature demonstrated that the primary requirements were knowledge and social support (Cavkaytar et al., 2014), and thus individuals with ID required support on daily life, social life, lifelong learn-

ing, employment, health and safety, and social adaptation (Thompson et al., 2004; Thompson et al., 2015). Therefore, PSIDA, which provided a social support resource for the parents by providing information on available social support mechanisms and how to plan for social support and provided information on the transition to adulthood, contributed to both the literature and the practice. Furthermore, it can be suggested that PSIDA sessions were social support mechanisms since they allowed mothers, who had similar psychosocial conditions, to share similar experiences, and to chat with each other, creating an environment of solidarity and sharing.

Previous studies demonstrated that family training programs, which aim to prepare the parents of individuals with ID for adulthood and the future, to reduce their stress, to improve their emotional well-being and self-efficacy, and to provide information and to support parenting and family interaction, reduced the stress levels of the parents, improved their social support, self-efficacy and quality of life perception levels, in other words, they had positive effects on various dimensions in the lives of the parents (Cankaya, 2013; Cavkaytar, 1999; Cavkaytar et al., 2014; Coren et al., 2018; Fodstad et al., 2018; Martinez-Rueda et al., 2017; Smith et al., 1996; Solomon et al., 1997; Yildiz & Cavkaytar, 2017). The parent training program developed in this study was found to be effective in changing the quality of life perception of parents. It is possible to say that this research finding coincides with the literature. It was determined that the parent training program developed in the present study was effective on changing the family quality of life perceptions of the participants. It is possible to argue that this study finding was consistent with the literature.

On the other hand, different findings were reported by applied family training studies conducted to improve the family quality of life perceptions. In a study conducted by Cavkaytar et al. (2014), it was observed that the family training program (E-FISTP) did not lead to a significant change in the quality of life perceptions of the parents, while leading to increases in their knowledge level, social support and their self-efficacy perceptions. In that study, the finding was explained by the facts that only one family member participated in the study, the limited use of the online environment by the parents, and the inadequacy of E-FISTP in changing quality of life perceptions since it was an information-oriented program, and it was stated that psycho-educational programs that provide more psychological support could lead to positive changes in quality of life perceptions. The fact that PSIDA improved the family quality of life perceptions in the present study, which was conducted with a single parent, was inconsistent with the above-mentioned finding in the literature. This discrepancy can be explained as follows: Since E-FISTP was developed based on the information requirements of parents on certain issues related to intellectual disability, the main objective was to provide information and support training for the parents. In contrast, in the present study, it can be suggested that PSIDA was effective in improving the quality of life since the program content included factors that affected the quality of life more such as adaptation to adulthood, preparing for the future, and coping with stress. As stated in the above-mentioned research, the fact that the main objective of E-FISTP was to provide information support and the fact that it was available online were not adequate to improve the quality of life perceptions, while the inclusion of available social support mechanisms and how to acquire these mechanisms, and inclusion of the activities that could be planned by the parents for their children and others and the fact that all sessions were conducted face to face and the activities were demonstrated with enactments in the present study were effective in improving the quality of life perceptions.

Martinez-Rueda et al. (2017) analyzed the effect of the family support program that they developed to improve the family quality of life perceptions using the Family Quality of Life Assessment tool and reported that the program was effective in improving the quality of life perceptions. In the present study, PSIDA improved the family quality of life perceptions. Thus, family training programs, which include content associated with psychological and social support that may affect the quality of life, could be effective in improving the family quality of life. Therefore, it can be suggested that the present study efficiency findings were consistent with the literature.

The present study findings demonstrated that PSIDA improved the quality of life perceptions of the families from low to moderate and led to a significant difference between experimental and control groups. However, review of the quality of life perceptions of the families on scale subdomains demonstrated that there were significant differences between the pre-test and post-test scores in two subdomains; an increase was observed in the perceptions of the parents in the experimental group in parenting and emotional well-being subdomains, while there were no differences between the pre-test and post-test scores of the same group in family interaction, physical well-being and disability support subdomain perception levels. In the literature, BCFQOL was used in one of the two studies that investigated the change in family quality of life perceptions after family training. In that study, as mentioned above, it was determined that the family training program did not lead to a change in family quality of life and subdomain perceptions (Cavkaytar et al., 2014). In contrast, in the present study, an increase was observed in the perception of two subdomains (parenting and emotional well-being) of family quality of life after family training.

The parenting subdomain includes concepts such as family members helping the children to stand on their own feet, teaching how to make decisions, teaching how to get along with others, helping children with their homework, and meeting their personal needs, while the subdomain of emotional well-being includes the support within the family to cope with stress, availability of peer support, having time to follow up their interests and presence of external support to meet special needs. In relation to these items, the fact that adaptation to adulthood, social support, and coping with stress were included in PSIDA content could have led to a change in parenting and emotional well-being perceptions. On the other hand, it was stipulated that other family members who live in the same house should also participate to observe a change in the subdomain of family interaction. Similarly, since the physical well-being subdomain reflects the economic status of the family, it is expected that economic status would not change with a training program. The subdomain of disability support includes the support provided by the school, family members, the environment and institutions. Since PSIDA is a program that provides information, guidance and support training for mothers, it would be normal not to affect the variables other than participants (schools, institutions, environment, etc.). Thus, the findings demonstrated that the program made a difference in the variables that could be influenced by the instructor who instructed the family training. Variables that were outside the control of the instructor were not affected by family training.

The participants of the study were determined based on voluntary participation and since only mothers volunteered to participate in the study, only mothers were included in the study. Coren et al. (2018) reported that the mothers were the participants in most studies in the literature on applied research on families. In this context, the study was consistent with the literature. This literature review demonstrated that information on the development of family training programs, their content, and the implementers of the program were only included in a small number of studies. In contrast, the present study included information on the determination of family needs, the development of the parent training program based on these needs, the content, the duration of the application and the instructor who implemented the program.

PSIDA was developed to reduce the stress and anxiety that the parents experience during the transition of their children with ID to adulthood and to support the adaptation to adulthood. The literature on applied studies on families with adult children demonstrated that there are limited studies that included the implementation of a program that prepares parents for adulthood. In this respect, it can be suggested that the present study is original.

The present study was carried out within the scope of the Assembly Work Vocational Education Program (AWVEP) project conducted at Eskisehir Tepebasi Municipality Ibrahim Ethem Kesikbas Assembly Workshop for Individuals with Disabilities with the support of ISKUR and Anadolu University. Since the project was a workshop where adults with ID are trained in a profession and acquire pre-

vocational and vocational skills before employment in sheltered workshops, it is one of the unique facilities available in Turkey. Thereby, the study was also unique in providing a parent training program on adult life and vocational education for the parents of the employees in the workshop.

The present research is limited to 18 participating mothers and PSIDA family training content. Furthermore, due to the fact that the study had to be completed in a short time and with intensive training within the scope of the project, the social validity of the study was not studied. However, the views of the mothers on the program were obtained at the end of the last session and these interviews were recorded. The views expressed in the video indicated that mothers were satisfied with the study, and that they were very upset that the program was over and demanded the continuation of the study. Another limitation of the study was the use of SIS-A in the stage of the determination of the needs, since no validity and reliability studies were conducted in Turkey.

The present study is an applied research on the families of adults with ID. In the literature, further studies on the transition process of individuals with ID to adulthood and support of the parents of these individuals in this process are required. Since this study was conducted with a small group, iteration studies are required. Further research could be planned with a higher number of participants in future studies, and the effect of family training programs on the level of knowledge of the parents could be analyzed. It may be advisable to plan longer-term studies on the family quality of life perceptions and to develop programs to improve these perceptions. Family training programs for adulthood in other disability groups could be developed and applied family training studies could be planned. Finally, all participants were mothers in the present study. The same study could be carried out with both parents or only with fathers.

References

Alwell, M., & Cobb, B. (2009). Functional life skills curricular interventions for youth with disabilities: A systematic review. *Career Development for Exceptional Individuals, 32,* 82–93.

American Association on Intellectual and Developmental Disabilities. (AAIDD). (2019). *AAIDD definition of intellectual disability.* https://aaidd.org/intellectual-disability/definition (Date accessed: 28.01.2019).

Arnett, J. J. (2010). *Adolescence and emerging adulthood.* (4th ed.). Pearson Education Limited.

Bailey, D. B., & Simeonsson, R. J. (1988). *Family assessment in early intervention.* Merrill Publishing Company.

Bayat, M., (2005). *How family members' perceptions of influences and causes of autism may predict assessment of their family quality of life.* Loyola Univesity Chicago. (Unpublished Doctoral Dissertation). UMI Microform 3180946.

Beach Center on Disability at the University of Kansas in partnership with families, service, providers and researhers (2006). *Family Quality of Life Survey.* Beach Center: University of Kansas.

Blacher, J. (2001). Transition to adulthood: Mental retardation, families, and culture. *American Journal on Mental Retardation, 106,* 173–188.

Brown, R. I., MacAdam–Crisp, J., Wang, M., & Iarocci, G. (2006). Family quality of life when there is a child with a developmental disability. *Journal of Policy and Practice in Intellectual Disabilities, 3,* 238–245.

Cankaya, S. (2013). *Zihin engellilere ozbakım ve ev ici becerilerinin ögretiminde ailelere yönelik beceri ögretimi yaziliminin gelistirilmesi ve degerlendirilmesi.* Yayımlanmıs Doktora Tezi. Eskisehir: Anadolu Universitesi.

Cavkaytar, A. (1999). Zihin engellilere ozbakım ve ev ici becerilerinin ögretiminde bir aile egitimi programının etkililigi. *Özel Egitim Dergisi, 2,* 40–50.

Cavkaytar, A., Ceyhan, E., Adiguzel, O. C., & Uysal, H. (2014). Effectiveness of online information and support services on the education of parents of children with intellectual disability. *DADD Online Journal,* 136.

Chou, Y. C., Lin, L. C., Chang, A. L., & Schalock, R. L. (2007). The quality of life of family caregivers of adults with intellectual disabilities in Taiwan. *Journal of Applied Research in Intellectual Disabilities, 20,* 200–210.

Coren, E., Ramsbotham, K., & Gschwandtner, M. (2018). Parent training interventions for parents with intellectual disability. *Cochrane Database of Systematic Reviews, 7.*

Davies, M. D., & Beamish, W. (2009). Transitions from school for young adults with intellectual disability: Parental perspectives on "life as an adjustment". *Journal of Intellectual and Developmental Disability, 34,* 248–257.

Drew, C. J., & Hardman, M. L. (2007). *Intellectual disabilities across the lifespan* (9th ed.). Pearson Merrill Prentice Hall.

Feldman M. A., Ducharme M., & Case L. (1999). Using self-instructional pictorial manuals to teach child-care skills to mothers with intellectual disabilities. *Behavior Modification, 23,* 480–497.

Friend, M. (2011). *Special Education: Contemporary perspectives for school professionals.* Pearson.

Fodstad, J. C., Kirsch, A., Faidley, M., & Bauer, N. (2018). Demonstration of parent training to address early self-injury in young children with intellectual and developmental delays. *Journal of Autism and Developmental Disorders, 48,* 1–12.

Gauthier-Boudreault, C., Gallagher, F., & Couture, M. (2017). Specific needs of families of young adults with profound intellectual disability during and after transition to adulthood: What are we missing? *Research in Developmental Disabilities, 66,* 16–26.

Grant, G., & Whittell, B. (2000). Differentiated coping strategies in families with children or adults with intellectual disabilities: The relevance of gender, family composition and the life span. *Journal of Applied Research in Intellectual Disabilities, 13,* 256–275.

Griffin, M. M., McMillan, E. D., & Hodapp, R. M. (2010). Family perspectives on post-secondary education for students with intellectual disabilities. *Education and Training in Autism and Developmental Disabilities, 45,* 339–346.

Hand, A., Raghallaigh, C. N., Cuppage, J., Coyle, S., & Sharry, J. (2012). A controlled clinical evaluation of the Parents Plus Children's Programme for parents of children aged 6–12 with mild intellectual disability in a school setting. *Clinical Child Psychology and Psychiatry, 18,* 536–555.

Hoffman, L., Marquis, J., Poston, D., Summers, J. A., & Turnbull, A. (2006). Assessing family outcomes: psychometric evaluation of the Beach Center Family Quality of Life Scale. *Journal of Marriage and Family, 68,* 1069–1083.

Hudson, A. M., Matthews, J. M., Gavidia-Payne, S. T., Cameron, C. A., Mildon, R. L., Radler, G. A., & Nankervis, K. L. (2003). Evaluation of an intervention system for parents of children with intellectual disability and challenging behaviour. *Journal of Intellectual Disability Research, 47,* 238–249.

Jokinen, N. S., & Brown, R. I. (2005). Family quality of life from the perspective of older parents. *Journal of Intellectual Disability Research, 49,* 789–793.

Keith, K. D., Heal, L. W., & Schalock, R. L. (1996). Cross-cultural measurement of critical quality of life concepts. *Journal of Intellectual and Developmental Disability, 21,* 273–293.

Kim, K. H., & Turnbull, A. P. (2004). Transition to adulthood for students with severe intellectual disabilities: Shifting toward person-family interdependent planning. *Research and Practise for Persons with Severe Disabilities, 29,* 53–57.

Knowles, C., Machalicek, W., & Van Norman, R. (2015). Parent education for adults with intellectual disability: A review and suggestions for future research. *Developmental Neurorehabilitation, 18,* 336–348.

Kraemer, B. R., & Blacher, J. (2001). Transition for young adults with severe mental retardation: School preparation, parent expectations, and family involvement. *Mental Retardation, 39,* 423–435.

Mactavish, J. B., MacKay, K. J., Iwasaki, Y., & Betteridge, D. (2007). Family caregivers of individuals with intellectual disability: Perspectives on life quality and the role of vacations. *Journal of Leisure Research, 39,* 127–155.

Martinez-Rueda, N., Fernandez-Gonzalez, A., Orcasitas-Garcia, J. R., Mortero-Centeno, D., & Villaescura-Peral, M. (2017). Construction of family quality of life support program for families of young and adults wtih intellectual disabilities. *Siglo Cero, 47,* 47–67.

McIntyre, L. L., Blacher, J., & Baker, B. L. (2002). Behaviour/mental health problems in young adults with intellectual disability: the impact on families. *Journal of Intellectual Disability Research, 46,* 239–249.

Meral, B. F. (2012). *Gelisimsel yetersizliği olan cocuklarin annelerinin aile yasam kalitesi algilarinin incelenmesi. (The Examination of the family quality of life perceptions of mothers who have children with disabilities).* Doctoral Thesis. Anadolu University, Eskisehir, Turkey.

Meral, B. F., & Cavkaytar, A. (2013). Beach Center Aile Yasam Kalitesi Ölçegi'nin Türkce uyarlama, gecerlik ve guvenirlik calismasi. *Egitim ve Bilim, 38,* 48–60.

Mugno, D., Ruta, L., D'Arrigo, V. G., & Mazzone, L. (2007). Impairment of quality of life in parents of children and adolescents with pervasive developmental disorder. *Health and Quality of Life Outcomes, 5,* 22.

Park, J., Hoffman, L., Marquis, J., Turnbull, A., Poston, D., Mannan, H., Wang, M., & Nelson, L. (2003) Assessing the family quality of life as the service outcome. *Journal of Intellectual Disability Research, 47,* 467–84.

Poston, D., Turnbull, A., Park, J., Mannan, H., Marquis, J., & Wang, M. (2003). Family quality of life: A qualitative inquiry. *Mental Retardation, 41,* 313–328.

Reseacrh Randomizer. (2016). Random sampling and random assignment. https://www.randomizer.org/ (Date Accessed: 12.12.2016)

Rossetti, Z., Lehr, D., Pelerin, D., Huang, S., & Lederer, L. (2016). Parent involvement in meaningful post-school experiences for young adults with IDD and pervasive support needs. *Intellectual and Developmental Disabilities, 54,* 260–272.

Schalock, R. L. (1994). Quality of life, quality en-

hancement, and quality assurance: Implications for program planning and evaluation in the field of mental retardation and developmental disabilities. *Evaluation and Training in Mental Retardation and Developmental Disabilities, 29,* 90–101.

Schalock, R. L., & Verdugo, M. A. (2002). *Quality of life for human service practitioners.* American Association on Mental Retardation.

Sileo, N. M., & Prater, M. A. T. (2011). *Working with families of children with special needs: Family and professional partnerships and roles.* Pearson Higher Ed.

Smith, G. C., Majeski, R. A., & McClenny, B. (1996). Psychoeducational support groups for aging parents: development and preliminary outcomes. *Mental Retardation, 34,* 172–188.

Smith, T. E. C., Gartin, B. C., Murdick, N. L., & Hilton, A. (2006). *Families and children with special needs professional and family partnerships.* Pearson Merrill Prentice Hall.

Solomon, P., Draine, J., Mannion, E., & Meisel, M. (1997). Effectiveness of two models of brief family education: Retention of gains by family members of adult with serious mentall illness. *American Journal of Orthopsychiatry, 67,* 177–186.

Steere, D.E., Rose, E. & Cavaiuolo, D. (2007). *Growing up transition to adult life for students with disabilities.* Pearson Education.

Stewart, D. (2009). Transition to adult services for young people with disabilities: current evidence to guide future research. *Develomental Medicine & Child Neurology, 51,* 169–173.

Sucuoğlu, B. (1995). Ozurlu çocugu olan anne-babalarin gereksinimlerinin belirlenmesi. *Cocuk ve Genclik Ruh Sagligi Dergisi, 2,* 10–18.

Thompson, J. R., Bryant, B., Campbell, E. M., Craig, E. M., Hughes, C., & Rotholtz, D. A., . . . Wehmeyer, M. L. (2004). *The Supports Intensity Scale (SIS): Users manual.* American Association on Mental Retardation.

Thompson, J. R., Bryant, B., Schalock, R. L., Shogren, K. A., Tasse, M. J., Wehmeyer, M. L. . . Rotholz, D. A. (2015). *Supports Intensity Scale – Adult version: Users manual.* American Association on Intellectual and Developmental Disabilities

Turnbull, A., Turnbull, H. R., Erwin, E., & Soodak, L. (2006). *Families, professionals, enad exceptionality. Positive outcomes through partnership and trust.* Merrill.

Turnbull, A., Turnbull, R., Wehmeyer, M. L., & Shogren, K. A. (2013). *Exceptional lives: Special education in today's schools.* Pearson.

Turnbull, A.P., Turbiville, V., & Turnbull, H.R. (2000). Evolution of family-professional partnership models: Collective empowerment as the model for the early 21st century. In J. P. Shonkof & S. L. Meisels (Eds.), *The Handbook of early childhood intervention* (2nd ed.; pp. 620–650). Cambridge University Press.

Vanderkerken, L., Heyvaert, M., Onghena, P., & Maes, B. (2018). Quality of life in flemish families with a child with an intellectual disability: A multilevel study on opinions of family members and the impact of family member and family characteristics. *Applied Research in Quality of Life, 13,* 779–802.

Viana Tomaz, R. V., Santos, V. D. A., da Silva de Avo, L. R., Ramos Germano, C. M., & Melo, D. G. (2017). Impact of moderate intellectual disability on the dynamics and quality of family life: A qualitative clinical study. *Cadernos De Saude Publica, 33.*

Walden, S., Pistrang, N., & Joyce, T. (2000). Parents of adults with intellectual disabilities: Quality of life and experiences of caring. *Journal of Applied Research in Intellectual Disabilities, 13,* 62–76.

Wehman, P. (2013). Transition: New horizons and challenges. In P. Wehman (Ed.) *Life beyond the classroom: Transition strategies for young people with disabilities* (pp. 3–41). Brookes Publishing Company.

Werner, S., Edwards, M., Baum, N., Brown, I., Brown, R. I., & Isaacs, B. J. (2009). Family quality of life among families with a member who has an intellectual disability: An exploratory examination of key domains and dimensions of the revised FQOL Survey. *Journal of Intellectual Disability Research, 53,* 501–511.

White, N., & Hastings, R. P. (2004). Social and professional support for parents of adolescents with severe intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 17,* 181–190.

Will, M. (1983). *OSERS programming for the transition of youth with disabilities: Bridges from school to working life.* Office of Special Education and Rehabilitative Services.

Yildiz, G., & Cavkaytar, A. (2017). Effectiveness of a sexual education program for mothers of young adults with intellectual disabilities on mothers' attitudes toward sexual education and the perception of social support. *Sexuality and Disability, 35,* 3–19.

Yoong, A., & Koritsas, S. (2012). The impact of caring for adults with intellectual disability on the quality of life of parents. *Journal of Intellectual Disability Research, 56,* 609–619.

Received: 13 February 2019
Initial Acceptance: 9 April 2019
Final Acceptance: 10 May 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 215–227

Evaluation of a Self-Instructional Package for Training ABA Service Providers to Conduct the Assessment of Basic Learning Abilities – Revised

Ashley L. Daniel, Karli Pedreira and Garry L. Martin
University of Manitoba

Toby Martin
University of Manitoba and St.Amant Research Centre

Abstract: The Assessment of Basic Learning Abilities-Revised (ABLA-R) is an assessment tool used by applied behavior analysis (ABA) service providers for determining the learning ability of clients with intellectual disabilities and autism spectrum disorder ASD. In this study, we combined an ABLA-R self-instructional manual (DeWiele et al., 2014) with video modeling and evaluated the effectiveness of the training package for teaching ABA service providers to administer the ABLA-R to the experimenter who role-played a child with ASD and to a child with ASD. We evaluated the training package using a single-subject, modified concurrent multiple-baseline (MB) design across a pair of participants, replicated across three other pairs of participants, plus a modified nonconcurrent MB design across a pair of participants, replicated across one other pair. Participants demonstrated a significant increase in the accuracy with which they administered the ABLA-R to an experimenter from baseline to post-test, and conducted the ABLA-R with high accuracy in a generalization assessment with a child with ASD. An implication of this research is that the self-instructional training package is an effective tool for training ABA service providers to accurately administer the ABLA-R to children with ASD.

The Assessment of Basic Learning Abilities-Revised (ABLA-R) test uses standard prompting and reinforcement procedures to assess the ease or difficulty with which a testee is able to learn a simple imitation and five two-choice discriminations. The ABLA-R is a valuable tool for matching the learning ability of individuals with intellectual disabilities (ID) and children with autism spectrum disorder (ASD) to the difficulty of training tasks. In this study we evaluated a self-instructional package for teaching ABA service providers to administer the ABLA-R to a child with ASD.

The original ABLA (prior to its revision) was developed by Kerr, Meyerson, and Flora (1977) and consisted of a motor imitation task (Level 1), a position discrimination task (Level 2), a visual discrimination task (Level 3), a visual-visual identity match-to-sample task (Level 4), a two-choice auditory discrimination task (Level 5), and an auditory-visual discrimination task (Level 6). Each task required a tester to use standardized prompting and reinforcement procedures to attempt to teach a client to perform the task until a standardized pass or fail criterion was met.

Research on the ABLA, reviewed by Martin et al. (2008) demonstrated that: (a) the ABLA tasks are ordered in difficulty as listed above; (b) the ABLA test has high test-retest and inter-tester reliability; (c) if a testee fails an ABLA level, the tasks requiring that level are very difficult to teach using standard prompting and reinforcement procedures; (d) mismatching of ABLA test level of clients to the ABLA difficulty of training tasks causes aber-

An expanded version of this paper was a doctoral dissertation submitted by the first author in partial fulfillment of the requirements of a Doctor of Philosophy degree in the Department of Psychology at the University of Manitoba. Correspondence concerning this article should be addressed to Ashley Daniel (adaniel@stamant.ca), or Garry Martin (Garry.Martin@umanitoba.ca). Mail correspondence should be sent to Garry Martin, Department of Psychology, University of Manitoba, 70 Dysart Road, 129 St. Paul's College, Winnipeg, MB, R3T 2M6, CANADA.

TABLE 1

A Brief Description of Levels of the Assessment of Basic Learning Abilities – Revised (ABLA-R)

Level	*Test Task*	*Everyday Examples*
Level 1 Imitation	When given a piece of foam, can the student imitate the teacher placing the foam into a container?	Children playing Follow-the-Leader. Rolling a ball from one person to the other.
Level 2 Position Discrimination	When presented with a yellow can and a red box in a stable position, can a student consistently place a piece of foam into the container on the left?	Turning on the cold (vs. the hot) water tap. Placing a fork on the left side of a plate when setting a table.
Level 3 Visual Discrimination	When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a piece of foam into the can?	Locating own printed name on the blackboard. Finding a particular shirt in a closet when the location changes each time it is replaced.
Level 4 Visual Identity Match-to-Sample Discrimination	When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a small yellow cylinder into the can, and a small red cube into the box?	Sorting socks into pairs. Restocking a partially emptied salad bar. Filling containers that are partly full.
Level 5 Visual Non-Identity Match-to-Sample Discrimination	When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a purple-colored piece of wood shaped like the word *CAN* into the can, and a piece of silver-colored wood shape like the word *BOX* into the box?	Placing a cup with a saucer. Putting a pencil with a piece of paper. Matching the printed word CAT to a picture of a cat.
Level 6 Auditory-Visual Combined Discrimination	When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a piece of foam into the correct container when the teacher requests either "red box" or "yellow can"?	Responding appropriately to the spoken words 'Stop' and 'Go'. Responding to requests such as 'Stand up' vs. 'sit down'.

Note: This table was adapted from DeWiele et al. (2011) with permission of the authors.

rant behaviours; (e) direct care staff with no knowledge of the ABLA test often mismatch the ABLA test level of clients and the ABLA difficulty level of training tasks; and (f) the ABLA test has concurrent validity with language, reading, and adaptive behaviour assessments.

Although some research demonstrated that the six original ABLA levels progressively increased in difficulty from Levels 1–6 (Martin et al., 1983; Vause et al., 2007), Martin and Yu (2000) reviewed several studies and concluded that most individuals who passed the original Level 5, an auditory discrimination task, were also able to pass the Level 6 task, an auditory-visual discrimination task. Since the original Level 5 task did not appear to serve as an intermediately difficult task between Levels 4 and 6, Sakko et al. (2004) evaluated the difficulty of a visual-visual nonidentity matching (VVNM) prototype task (see Table 1). They found that the VVNM task was an appro-

priate level of difficulty to serve as a Level 5 task, with high test-retest reliability and high predictive validity for an individual's ability to learn other VVNM tasks. DeWiele et al. (2011) replaced the original Level 5 task with Sakko et al.'s VVNM task and named the improved test the ABLA-R. Research findings on the ABLA-R are consistent with previous research on the ABLA (Yu et al., 2015).

The ABLA-R Testing Materials and Procedure

The materials that are required to administer the ABLA-R are identical to those required for the original ABLA, with the exception of the materials required to test the new Level 5 task. The materials include: a small piece of white foam, a small yellow cylinder, a small red cube, a large yellow can, and a large red striped box, small edible reinforcers, and the ABLA-R datasheets. For Level 5, the required materials also include the word *CAN* made from a purple material, and the word *BOX* made from a silver material.

For each level of the ABLA-R, the target behavior of the testee is to place the manipulandum (i.e., the foam, red cube, yellow cylinder, or the word *BOX* or *CAN*) into the correct container (i.e., the red box or yellow can), which will require the testee to imitate the tester (Level 1), or discriminate between the correct container and a distractor container (Levels 2–6). The tester begins testing a level with an initial prompting sequence that consists of a demonstration of the target response, a prompted trial where the testee performs the response, and then an opportunity for the testee to perform the response independently. After the testee performs an independent response, the tester would proceed with administering the test trials for the level. At the start of each test trial, the tester gives a manipulandum to the testee, and presents a verbal instruction. For Level 1, which tests an individual's ability to perform motor imitation, and Levels 2–5, which assess an individual's ability to perform different visual discriminations, the instruction, "Where does it go?" prompts the testee to place the manipulandum into the correct container. For Level 6, which assesses an individual's ability to perform an auditory-visual discrimination, the tester indicates the correct container by using a high-pitched tone of voice to quickly say, "red box," or a low-pitched tone to slowly say, "yellow can." Each correct independent response is reinforced with praise and a small edible. Each incorrect response is followed by a verbal statement (e.g., "No, that's not where it goes") and an error correction trial that follows the same procedure as the initial prompting sequence. The test trials continue until a testee has performed the target behavior correctly on eight consecutive independent trials, which is considered a "pass," or until a testee has performed eight cumulative errors during the initial prompting sequence and/or test trials, which is considered a "fail." Typically, a testee will pass or fail a level within 30 trials (DeWiele et al., 2011).

Training Individuals to Implement the ABLA-R

Since the late 1990s, researchers have developed and evaluated several instructional training tools for effectively and efficiently teaching service providers to administer the original ABLA and the ABLA-R. To help the reader follow the history of research on these instructional tools, an overview is provided in Table 2. The original information package provided by Kerr et al. (1977, see Table 2) included a description of the test procedure that was published in a monograph issue on the original ABLA in *Rehabilitation Psychology*, an informational summary of results from previous studies on the ABLA, and three informational appendices. Kerr et al. did not evaluate the effectiveness of that information package. In 1998, DeWiele and Martin prepared a self-instructional manual (ABLA-SIM_a, see Table 2). DeWiele et al. (2000) conducted an experiment to assess the effectiveness of the ABLA-SIM_a for teaching a sample of 21 undergraduate university students how to administer the original ABLA. DeWiele et al. compared the self-instructional manual to the best alternative training material available at the time, which was the information package by Kerr et al. (1977), plus a brief written introduction to the ABLA. The experiment included six dependent variable measures of ABLA knowledge. Participants were randomly assigned to study either the self-instructional manual or

TABLE 2

An Overview of Versions of the ABLA and ABLA-R, and Instructional Manuals for Each Version

Version of the ABLA	*ABLA Levels*	*Instructional Manuals*
Original ABLA	1, 2, 3, 4, original 5, & 6	Information package (Kerr et al., 1977) ABLA Self-Instructional Manual (ABLA SIM_a; Wiele & Martin, 1998)
Shortened ABLA	1, 2, 3, 4, & 6	ABLA SIM_b (DeWiele et al., 2000)
ABLA-R	1, 2, 3, 4, new 5, & 6	ABLA-R SIM (DeWiele et al., 2011) ABLA-R SIM-R (DeWiele et al., 2014)

the information package, and then completed the dependent measure assessments. The results indicated that, on average, the group that studied the self-instructional manual obtained a higher score on all six dependent measures than the group that studied the information package. In the role-play assessment with a confederate, participants' accuracy of administering the ABLA was scored using a procedural checklist. Participants in the self-instructional manual group administered the original ABLA with a mean accuracy of 85.6%, and participants in the information package group administered the test with a mean accuracy of only 68.1%, which was a statistically significant difference.

ABLA SIM_b

DeWiele et al. (2000) then made several revisions to simplify the information in the manual; the most notable revision was removing Level 5 from the ABLA procedure, based on the data reviewed previously that most individuals who passed Level 5 also passed Level 6. This produced a shortened ABLA, which led to ABLA SIM_b (DeWiele et al., 2000, see Table 2). The shorter manual was tested with four direct-care staff. The participants studied the ABLA SIM_b for a mean of 5 hours and 55 minutes, after which they completed written exams for comprehension, fluency, and task classification. DeWiele et al. (2000) required participants to re-write the exams until a 90% mastery criterion was achieved on each exam. After participants had mastered the three different exams, they each conducted an applied ABLA assessment with three individuals with a DD. Although participants' baseline ability to implement the shortened ABLA was not assessed, they did not have prior experience with the ABLA. The mean procedural accuracy of the direct-care staff's applied assessments with individuals with DD was 82%. The results of this study indicated that direct-care staff administered the shortened ABLA with high accuracy after studying the ABLA SIM_b, although the mean procedural accuracy obtained by participants in this experiment was lower than in the previous experiment testing the original ABLA by DeWiele et al. (2000).

ABLA-R and ABLA-R SIM

As described previously, research reviewed by Martin and Yu (2000) revealed that a high percentage of individuals who pass the original Level 5 of the ABLA also typically pass Level 6, and research by Sakko et al. (2004) demonstrated that a visual-visual nonidentity matching (VVNM) prototype task was a suitable task for assessing a discrimination skill that is less difficult than Level 6 but more difficult than Level 4. Therefore, the ABLA self-instructional manual was further revised to include the VVNM prototype task as a new Level 5 task (DeWiele et al., 2011, see Table 2). The version of the ABLA that includes the VVNM Level 5 task is called the ABLA-R, and the self-instructional manual for the ABLA-R is called the ABLA-R SIM (DeWiele et al., 2011).

Prior to assessing the effectiveness of the ABLA-R SIM, Martin, Martin, Yu, Thomson, Boris, and DeWiele (2014) developed the *Assessment of Basic Learning Abilities-Revised Tester Evaluation Form.* Research indicates that the ABLA-R TEF is a valid and reliable tool for

assessing the quality of an ABLA-R assessment (Awadalla et al., 2014). Based on that research, the ABLA-R TEF was revised and is now the ABLA-R TEF-R (Martin et al., 2015).

Boris et al. (2015) assessed the ABLA-R SIM to teach eight university students to administer the ABLA-R. They used a modified multiple-baseline design across a pair of participants that was replicated across four pairs. In a baseline and post-test assessment, participants administered three levels of the ABLA-R to a confederate who role-played an individual with DD. Participants' accuracy of administering the ABLA-R was measured using the ABLA-R Tester Evaluation Form (ABLA-R TEF-R; Martin et al., 2015). During training, participants independently studied the ABLA-R SIM, although they were prompted to complete the study exercises in the manual. Participants were required to achieve 100% accuracy on a written mastery test containing study questions from the manual prior to completing the post-training assessments. On average, participants spent 2 hours and 43 minutes during training, but this did not include the amount of time required for participants to write the mastery test. Across participants, the mean improvement in ABLA-R procedural accuracy from baseline to post-training assessments was 60%, and participants' mean post-training, as measured by the ABLA-R TEF-R, accuracy was 83%. Six out of eight participants achieved at least 90% accuracy in the post-training assessment and administered the ABLA-R to an individual with DD as a measure of generalization, with a mean accuracy equaling 91%. The results of this study indicated that the ABLA-R SIM was effective for teaching most of the participants to administer the ABLA-R with 90% or higher accuracy. However, Boris et al. also examined the mean accuracy with which participants performed each item of the ABLA-R TEF during a post-training assessment and found that participants performed some items with low accuracy (defined as less than 80% accuracy).

In the current research, the limitations of the study by Boris et al. (2015) were addressed in three ways. First, with the support of the authors of the ABLA-R SIM, the manual was revised to more effectively teach all items of the ABLA-R testing procedures, including the items that were performed with low accuracy. The revised ABLA-R SIM is referred to as the ABLA-R SIM-Revised (ABLA-R SIM-R; DeWiele et al., 2014). Second, six training videos, one for each level of the ABLA-R were combined with the ABLA-R SIM-R to produce a training package. Third, the new training package was implemented with ABA service providers to teach them how to administer the ABLA-R to a child with ASD.

Method

Ethical approval of the protocol was obtained from the University of Manitoba Psychology/Sociology research Ethics Board and the St. Amant Decision of Research Access Committee prior to conducting the study.

Participants and Setting

ABA service providers. The participants in this study included 12 staff members employed by St.Amant, a government-funded ABA program for children with ASD located in Winnipeg, Manitoba. The service providers were recruited through mailed recruitment packages and an e-mail sent by the ABA program manager, and a recruitment presentation. The Informed Consent Form clearly communicated to the service providers that agreeing or declining to participate in the study would not affect the status of their employment in the program. Autism senior tutors and autism consultants were required to receive training on the ABLA-R as part of their employment. However, they could become trained via an alternative training procedure in the workplace if they did not wish to participate in the study. Autism tutors (often referred to as "therapists" or "technicians" by other agencies) were not required by their employer to receive training on the ABLA-R. All staff participants received an honorarium of $40.00, and the autism senior tutors and autism consultants were also compensated by their employer for up to six hours of training.

Among the 12 ABA service providers, there were five autism tutors, two autism senior tutors, and five autism consultants. One out of the five consultants (Participant 6) was a

TABLE 3

ABA Service Provider Characteristics

No.	*Position*	*Age Range*	*Gender*	*Experience in Current Position*	*First Language is English*	*Education*	*Major*
P1	Tutor	36-40	F	1–2 years	Yes	Bachelor's degree	Recreation Studies
P2	Consultant	26–30	F	6 months–1 year	Yes	Master's degree in progress	Psychology
P3	Consultant	26–30	F	< 6 months	No	Master's degree	Behaviour Analysis
P4	Consultant	21–25	F	6 months–1 year	Yes	Master's degree in progress	Psychology
P5	Consultant	26–30	F	< 6 months	No	Master's degree	Biology/Psychology
P6	Consultant	31–35	M	< 6 months	Yes	Master's degree	Behaviour Analysis
P7	Tutor	21–25	F	< 6 months	Yes	Postsecondary	N/A
P8	Senior Tutor	26–30	F	< 6 months	Yes	Bachelor's degree	Psychology
P9	Tutor	21–25	F	< 6 months	No	Bachelor's degree	Psychology
P10	Senior Tutor	26–30	F	1–2 years	Yes	Bachelor's degree	Psychology
P11	Tutor	21–25	F	6 months–1 year	Yes	Bachelor's degree in progress	Psychology
P12	Tutor	21–25	F	6 months–1 year	No	Bachelor's degree	Psychology

board certified behaviour analyst (BCBA). Nine participants did not have any prior training on the ABLA-R. Out of the three participants who indicated they had received prior training, Participant 2 had read a previous version of the training manual but had not been exposed to the ABLA-R for one year, Participant 4 had briefly read a previous version of the training manual and observed another person conducting the assessment, and Participant 8 had read part of the manual and observed another person conducting the assessment. Other characteristics of the staff participants are displayed in Table 3.

Participant with ASD. One child with ASD participated in the generalization phase of this study. The child was recruited through a school age program via a recruitment package mailed to his parents by St.Amant's privacy officer. It was clearly indicated that accepting or declining to participate in the study would not affect the services that they received from the program. The child was an 11-year-old male who had been diagnosed with ASD, and had received over seven years of ABA services. Prior to conducting generalization sessions with the ABA service providers, the experimenter conducted the ABLA-R with the child and concluded that he was able to pass all six levels.

Setting. All sessions occurred in a testing room at a residential and treatment center for individuals with ID. The testing room contained a table and chairs for the participants, a telephone, and a cabinet for storage, but was otherwise unfurnished.

Materials

In the baseline, training, post-test, and generalization assessments, the ABLA-R TEF-R was used to measure the accuracy with which participants conducted the ABLA-R. A laptop was used to video-record the sessions in all phases of the study.

In baseline, participants were given the same brief instructions which outlined the procedure for administering the ABLA-R that was used by Boris et al. (2015); the ABLA-R testing materials, edible reinforcers; and the ABLA-R datasheets. The experimenter role-played a client and followed a script for responding during baseline, which indicated whether to respond correctly or incorrectly on a trial.

During training, the ABLA-R testing materials were provided, as well as a self-instructional training package that included the ABLA-R SIM-R and a video modeling component.

ABLA-R SIM-R. The training package contained the ABLA-R SIM-R that was revised (DeWiele et al., 2014) based on the results of the study by Boris et al. (2015). Seventeen items in total were performed with low accuracy in Boris et al. (2015). The information pertaining to those items was expanded upon in the ABLA-R SIM-R.

Other revisions to the manual based on Boris et al. (2015) included simplifying the Summary of Steps for testing each level, inserting the Summary of Steps for each level after its testing description, adding a role-play exercise for Level 5, inserting a sample of a scored Level 5 data sheet, and adding descriptions of the low accuracy items to the "Frequent Errors that Testers Make" section.

Video modeling. The training package also contained six video demonstrations; each video corresponded to a level of the ABLA-R. The average duration of a video was 3 minutes and 37 seconds (range: 2m 33s – 4m 20s). Each video began by showing the materials that are needed to conduct a particular level. In each video, an experimenter demonstrated conducting the ABLA-R with a confederate learner. The following scenarios were modeled: (a) the initial prompting sequence for a level; (b) a test trial on which the testee responded correctly and the tester provides the consequence for a correct response; (c) a test trial on which the testee responded incorrectly, followed by the tester administering an error correction; and (c) several trials in which the testee responds correctly and incorrectly, with a prompt for the participant to practice scoring the confederate's responses in the video. The videos included narrative vocal prompts to increase the likelihood that the participants would attend to the demonstrations. After the scoring practice section, a correctly completed datasheet was shown so that a participant could compare his/her scoring.

Post-test and generalization. The materials required for the post-test and the generalization assessment included the ABLA-R testing materials, data sheets, edible reinforcers, and a list of steps from the manual that summarized how to test each level. In the post-test, the experimenter followed the same confederate scripts for each level that were used in baseline. The reinforcers used in the generalization assessment were edibles provided by the child's parents or the experimenter (e.g., gummy bears).

Experimental Design and Dependent Variable

The experimental design used in this study was a single-subject, modified concurrent multiple baseline (MB) design across a pair of participants (Martin & Pear, 2015). This design was replicated across three additional pairs of participants for a total of eight participants. Due to difficulty recruiting participants and scheduling concurrent sessions, the experimenter also used a modified nonconcurrent MB design across a pair of participants that was replicated across one other pair. The modification to the MB designs included fewer baseline sessions than a typical MB design. The participants who were evaluated using modified concurrent MB designs were Participants 3 and 4, Participants 5 and 6, Participants 9 and 10, and Participants 11 and 12. Participants 1 and 2, and Participants 7 and 8 were assessed using modified nonconcurrent MB designs. Participants were assigned to pairs based on their availability and the order in which they volunteered to participate in the study.

The length of the baseline phase was predetermined for all participants. In each MB design across participants, the baseline performance of one participant (shown on the bottom panel of each graph) was assessed on two separate occasions, and the participant conducted a total of six ABLA-R levels. The baseline performance of the other participant (shown on the top panel of each graph) was assessed one time on three of the levels.

The independent variable in this study was exposure to the self-instructional training package, and the dependent variable was participants' ABLA-R accuracy. Accuracy was measured as percentage correct on the ABLA-R TEF, and was defined as the number of steps that a participant performed correctly divided by the total number of steps required to correctly conduct a level, multiplied by 100%.

Procedure

The study involved four phases: (a) a baseline assessment; (b) implementation of a self-in-

structional training package; (c) a post-test; and (d) a generalization assessment.

Baseline. Participants were given a time limit of 10 minutes to study brief instructions and a data sheet for conducting each predetermined level of the ABLA-R, but the instructions were available for the participant to reference throughout the entire assessment. After studying the instructions, a participant then attempted to administer three levels of the ABLA-R to an experimenter, who role-played a child with ASD. The experimenter's response script, previously used by Boris et al. (2015), ensured that her behavior was standardized across participants.

The levels on which a participant was assessed in baseline were selected in a manner that ensured the participant was assessed on a variety of different levels. Although the steps that a tester is required to perform varies across ABLA-R levels, there are some levels for which the procedure is very similar. Therefore, in order to assess a participant on his/her ability to conduct a variety of levels, a participant was not assessed on both Levels 2 and 3, or Levels 4 and 5 in the same session. For example, if Level 2 was randomly selected, and the next level selected was Level 3, this selection would be discarded and the random number generator would select another level for the assessment. Once a level was selected, it could not be selected again. The first participant in each pair was assessed on three levels in baseline (in one session), and the second participant was assessed on all six levels in baseline (in two sessions). The order in which levels were tested varied, based on the order in which they were selected. Table 4 shows the levels on which participants were assessed in baseline (and post-test), in the sequence in which they were assessed.

Self-instructional training package. A participant was provided with the ABLA-R SIM-R, a laptop loaded with the six demonstration videos, ABLA-R datasheets, and all of the materials required for conducting the ABLA-R.

The manual prompted the participant to watch the corresponding demonstration video for a level immediately after the description of the procedure. Additionally, the manual instructed a participant to practice testing each level with a partner to play the role of a client. The manual provided a script to guide the partner's responses. The experimenter served as the participant's partner during the role-play exercises; however, no feedback was given to a participant regarding his/her performance in order to maintain the integrity of the self-instructional training package. The role-play exercises were videotaped and scored at ausb later time to measure a participant's progress throughout the training phase.

TABLE 4

Levels Assessed in Baseline and Post-Test

	Levels Assessed	
Participant	*Baseline*	*Post-Test*
P1	4, 1, 3	4, 1, 3
P2	6, 1, 5, 2, 4, 3	6, 1, 5
P3	1, 3, 4	1, 3, 4
P4	6, 3, 5, 4, 2, 1	6, 3, 5
P5	2, 3, 4	2, 3, 4
P6	6, 1, 4, 3, 5, 2	6, 1, 4
P7	3, 5, 1	3, 5, 1
P8	1, 4, 5, 6, 2, 3	1, 4, 5
P9	1, 5, 4	1, 5, 4
P10	3, 5, 6, 2, 1, 4	3, 5, 6
P11	4, 6, 5	4, 6, 5
P12	6, 3, 2, 1, 4, 5	1, 4, 5

Post-test. Prior to conducting the post-test, a participant was asked to complete a social validity questionnaire. For the post-test, a participant was provided with the ABLA-R testing materials and data sheets that were provided in baseline. A participant was informed that he/she would be asked to conduct three levels of the ABLA-R, and was given a summary of steps from the manual for testing a specific level of the ABLA-R that he/she could refer to if needed during the assessment. The three levels on which a participant was assessed during the post-test were the first three assessed in baseline, in the same order. The exception was Participant 12, who was assessed on the last three levels assessed in baseline.

The experimenter reminded the participant that she would play the role of the client and would not be able to provide feedback on the participant's performance. Once a participant indicated that he/she was ready to begin the assessment, the experimenter said, "Now please try to assess the researcher. Act as if you were actually assessing a child with autism."

After a participant had completed assessing the experimenter on a level, the participant was provided with a summary of steps for testing another level of the ABLA-R, and was asked to assess the experimenter on that level. This procedure was repeated for the third level of the post-test.

Generalization. The external validity of this study was addressed by measuring the accuracy with which participants administered the ABLA-R to a child with ASD. Only participants who achieved a mastery criterion of at least 80% on each level with an overall mean of at least 90% accuracy across the three levels in the post-test were assessed in generalization. Generalization occurred, on average, 63 days after a participant's post-test. However, the length of time between post-test and generalization was variable across participants (range: 40–88 days). The experimenter informed the staff participant that he/she could refer to the summary of steps during the assessment, and that she could not provide any feedback to the participant. The experimenter stated, "You will begin at Level 1 and continue until you think you should stop." Once the staff participant indicated that he/she had completed the testing, the session was ended.

Interobserver Agreement (IOA)

Two trained observers independently scored 35% of the levels from video (a total of 59 levels) using the ABLA-R TEF. An agreement occurred when the observers both scored an item as correct or both scored an item as incorrect. A disagreement occurred when the observers did not score an item identically. Percent agreement was computed for each observed session by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2015). The average percent agreement across all phases and sessions was 89.71% (range: 47.69%–100%).

Procedural Integrity

The experimenter scored her own behaviour using a procedural integrity checklist for each phase. If the experimenter followed a step as planned, she recorded a correct response for that item. If she missed a step or did not follow the procedure as planned, she recorded an error for that item. The procedural integrity was calculated by dividing the number of items that the experimenter followed correctly by the total number of items and multiplying by 100%. The procedural integrity was 100%.

Results

Participants' accuracy in administering the ABLA-R is illustrated in Figures 1 and 2. The average accuracy with which participants administered the ABLA-R was 43.0% in the baseline assessment (range: 21.3%–63.5%), 92.3% in the training phase (range: 85.0%–98.2%), and 92.7% in the post-test (range: 74.2%–99.5%). The average improvement in accuracy from baseline to post-test was 49.7%. Eight participants (P1, P2, P4, P5, P6, P9, P10, and P12) performed at mastery level in the post-test and thus qualified to participate in generalization phase. One participant (P12) who met the mastery criteria did not respond to an invitation to participate in the generalization assessment. The average ABLA-R accuracy of the seven participants who conducted the ABLA-R with a child with ASD in the generalization assessment was 93.7% (range: 87.8%–96.2%).

Visual inspection revealed that the stability of baseline data was variable across participants. Baseline accuracy scores typically fell near or below 50%, however some participants demonstrated higher accuracy in baseline. Nevertheless, visual inspection revealed no overlapping data points between baseline and post-test for 11 out of 12 participants, a sizeable improvement from baseline to post-test was observed for the majority of participants, and replication of an effect across all participants.

Participants correctly ended the testing of a level 77.78% of the time during the post-test, and 85.71% of the time during generalization. However, all participants in the generalization phase correctly assessed the child's learning ability at Level 6; that is, the participants terminated the assessment only after the child passed all six levels. Therefore, even though the participants in generalization made errors in determining

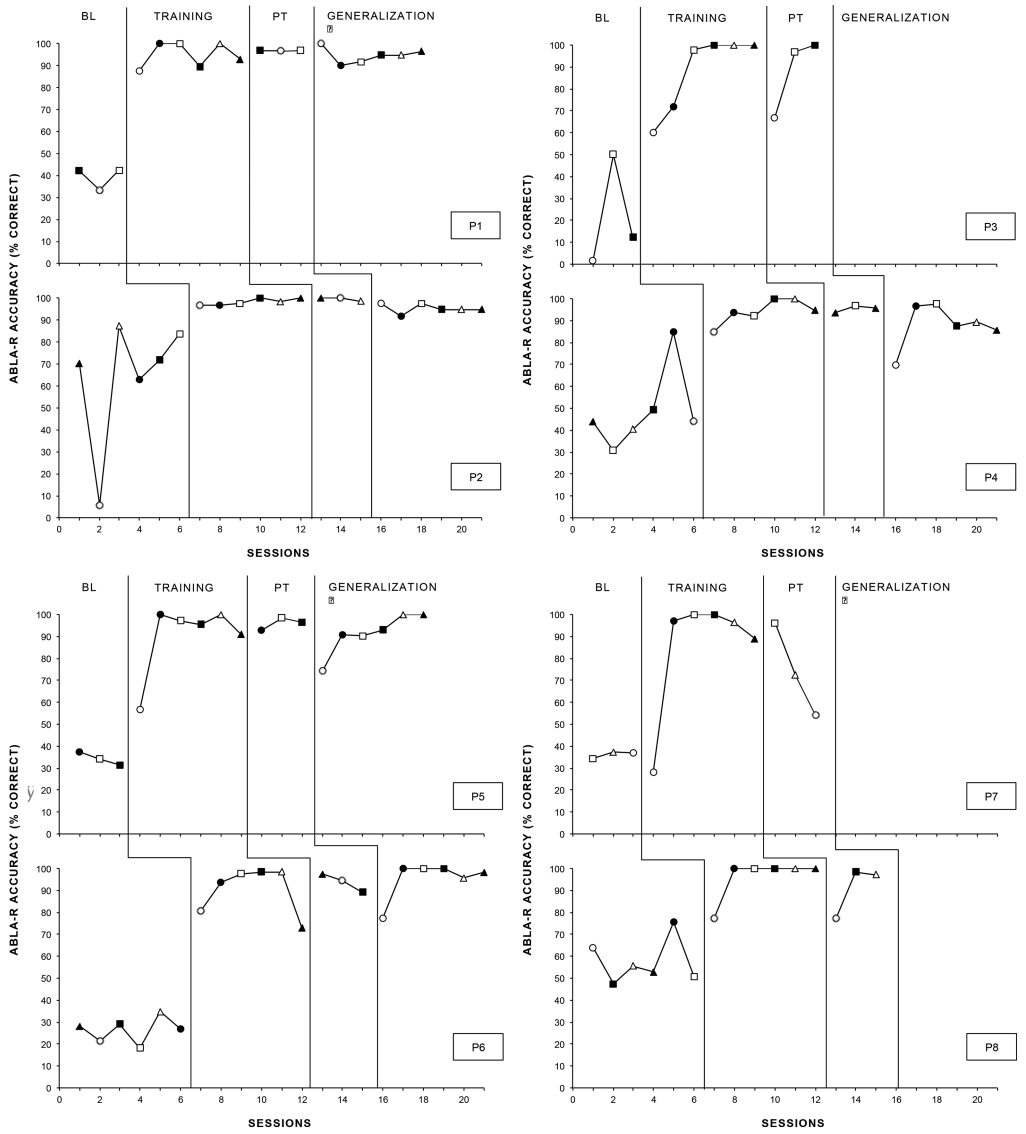


Figure 1. Accuracy in administering the ABLA-R, shown as percent correct on the ABLA-R TEF for Participants 1–8. ABLA-R levels are represented as follows: Level 1 (○), Level 2 (●), Level 3 (□), Level 4 (■), Level 5 (△), and Level 6 (▲).

when to stop testing a level, they all arrived at the correct outcome.

The average amount of time that participants needed to complete the self-instructional training package was 2 hours and 23 minutes (range: 1h 45m – 3h 35m). This included time spent reading the manual, watching the demonstration videos, and engaging in the role-play exercises with the experimenter serving as the client.

Social Validity

All participants completed a 7-item social validity questionnaire after training. Overall, participants either rated the ABLA-R SIM-R as very easy to read ($N = 7$), or moderately easy to read ($N = 5$). Participants rated the study questions in the manual as either very useful in helping them to learn about the ABLA-R

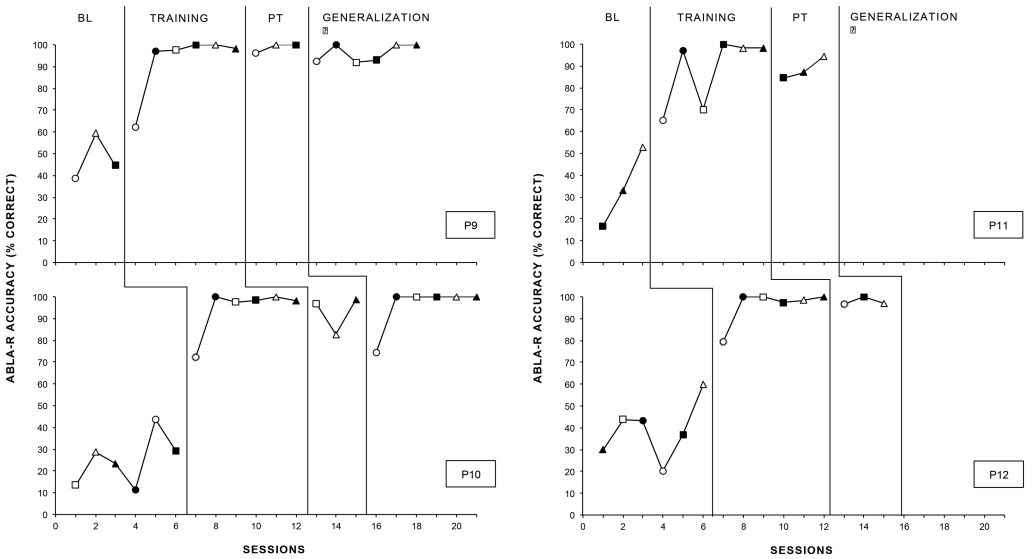


Figure 2. Accuracy in administering the ABLA-R, shown as percent correct on the ABLA-R TEF for Participants 9–12. ABLA-R levels are represented as follows: Level 1 (○), Level 2 (●), Level 3 (□), Level 4 (■), Level 5 (△), and Level 6 (▲).

($N = 8$), or moderately useful ($N = 5$). When participants were asked to rate the helpfulness of the ABLA-R SIM-R overall in preparing them to conduct the ABLA-R, they indicated that it was either very useful ($N = 6$), or moderately useful ($N = 6$). All participants indicated that they watched all of the demonstration videos, and most indicated that they were very attentive when watching the videos (N=11), while one participant was moderately attentive. All participants rated the demonstration videos as being very helpful in preparing them to conduct the ABLA-R.

Discussion

It was hypothesized that the self-instructional training package would increase the accuracy with which ABA service providers would be able to conduct the ABLA-R. An immediate increase in accuracy was observed for 11 of the 12 participants upon introduction of the ABLA-R SIM-R and video modeling during training, and 8 out of 12 participants achieved a high level of mastery when conducting the ABLA-R with the experimenter in the post-test. A calculation of effect size between baseline and post-test also revealed a large effect. In the generalization phase, seven participants conducted the ABLA-R with a child with ASD with high accuracy across all six levels, and correctly determined the child's ABLA-R level. Overall, the results suggested that participants successfully learned to conduct the ABLA-R using the self-instructional training package.

These results are consistent with the results found by Boris et al. (2015), and extend that study by demonstrating the generalizability of the ABLA-R SIM-R plus video modeling for training a sample of participants from the target population of ABA service providers. This study also expanded on Boris et al. by assessing the ability of trained participants to conduct the ABLA-R with a child with ASD. However, the average improvement in accuracy from baseline to post-test was slightly less in the current study than in Boris et al. A potential explanation would be that since some of the ABA service providers in this study had prior knowledge or exposure to the ABLA-R, they performed at a higher level in the baseline phase than the students who participated in the Boris et al., and those with higher accuracy in baseline had less room for improvement.

The participants completed the training package in a relatively short amount of time. The average amount of time required in the current study was 2 hours and 23 minutes, which was comparable to the 2 hours and 43 minutes that participants required in Boris et al.'s (2015) study.

The training package was shown to be socially valid according to the participants, who provided positive ratings for both the ABLA-R SIM-R and the demonstration videos. Participants indicated that the manual was easy to read, that the study questions in the ABLA-R SIM-R were helpful, and that the video modeling component was very helpful in teaching them to conduct the ABLA-R. The social validity of an intervention is an important factor in determining whether or not it will be accepted and implemented by consumers (Cooper, Heron, & Heward, 2007). While consumer satisfaction is not a substitute for the empirical support of an intervention, Hawkins (1991) indicated that it is a powerful influence on the opinion of decision-makers, such as funding sources and boards of directors, who are more likely to support a socially valid intervention or training tool.

The results should be interpreted with caution, as there are some limitations to this study. First, three of the participants had previously been exposed to the ABLA-R and/or ABLA-R training, and this was evident in their high baseline scores. Some participants also demonstrated an increasing trend in baseline, or variable baseline scores. An increasing trend may have occurred due to practice effects, given the similarity in the testing procedure of some levels. An extended baseline phase would have clarified whether the baseline accuracy for those participants would have stabilized or continued to increase without training. However, due to the limited time available with the staff participants, and the urgency with which the staff needed to be trained, it was not possible to extend the baseline phase. Second, participants' performance of Level 1 was typically less accurate than performance of the other levels during the post-test and/or generalization. Boris et al. (2015) also observed this pattern with the participants in their study. The revisions to the ABLA-R SIM-R, as well as the demonstration videos, attempted to correct this issue by clarifying the procedure for Level 1, however it appears that some participants still found this level challenging in comparison to the other levels. In the future, researchers could attempt to further improve training methods for Level 1, perhaps by providing extra video modeling or by incorporating an additional role-play exercise for Level 1. Third, there are some limitations regarding the generalization assessment. Although participants demonstrated high accuracy when conducting the ABLA-R with a child with ASD, this assessment was conducted approximately two months after the post-test, and it is unknown whether participants had been exposed to the ABLA-R or conducted the ABLA-R with their clients in the interim.

In conclusion, the findings of the current study indicate that combining the ABLA-R SIM-R plus video modeling resulted in an effective training package for increasing the accuracy with which ABA service providers conducted the ABLA-R with the experimenter role-playing a client. The ABA service providers who performed at mastery level with the experimenter also generalized their assessment skills to a child with ASD. It is important for ABA service providers to accurately conduct the ABLA-R as this assessment can provide essential information about a client's learning ability, which can be used to tailor an individual's behavioural programming. This research contributes to the larger area of staff training by providing support for self-instructional training methods that involve written instructions and video modeling. The practical significance of these findings is that employers can utilize self-instructional training for teaching staff to conduct behavioural assessments and procedures in a cost-effective and time-efficient manner.

References

Awadalla, N., Boris, A. L., Wightman, J. K., Miljkovic, M., Kaminski, L., Martin, T. L., Martin, G. L., & Yu, C. T. (2014). Evaluation of a tester evaluation form for the Assessment of Basic Learning Abilities – Revised. *Journal on Developmental Disabilities, 20,* 91–101.

Boris, A. L., Awadalla, N., Martin, T. L., & Martin, G. L. (2015). An evaluation of a self-instructional manual for teaching individuals to administer the

revised ABLA test to persons with intellectual disabilities. *Education and Training in Autism and Developmental Disabilities, 50,* 466–479.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

DeWiele, L. A., & Martin, G. L. (1998). *The Kerr-Meyerson Assessment of Basic Learning Abilities: A self-instructional manual.* Unpublished manuscript, available from G. Martin, Psychology Department, University of Manitoba, Winnipeg, Manitoba, Canada, R3T 2M6.

DeWiele, L., Martin, G. L., & Garinger, J. (2000). Field testing of a self-instructional manual for the ABLA test. *Journal on Developmental Disabilities, 7,* 93–108.

DeWiele, L., Martin, G. L., Martin, T., Yu, D., & Thomson, K. (2011). *The Kerr-Meyerson Assessment of Basic Learning Abilities Revised: A self-instructional manual.* Retrieved from: http://stamant.ca/research/abla

DeWiele, L., Martin, G. L., Martin, T., Yu, D., Thomson, K, & Boris, A. L. (2014). *The Kerr-Meyerson Assessment of Basic Learning Abilities Revised: A self-instructional manual - revised.* Retrieved from: http://stamant.ca/research/abla

Hawkins, R. O. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis, 24,* 205–213.

Kerr, N., Meyerson, L., & Flora, J. (1977). The measurement of motor, visual, and auditory discrimination skills. *Rehabilitation Psychology, 24,* 95–112.

Martin, G., Martin, T., Yu, D., Thomson, K., Boris, A., & DeWiele, L. (2014). *ABLA-R Tester Evaluation Form.* Retrieved from: http://stamant.ca/research/abla

Martin, G., Martin, T., Yu, D., Thomson, K., Boris, A., & DeWiele, L. (2015). *ABLA-R Tester Evaluation Form - Revised.* Unpublished manuscript.

Martin, G. L., & Pear, J. J. (2015). *Behavior modification: What it is and how to do it* (10th ed.). Pearson Education.

Martin, G. L., Thorsteinsson, J. R., Yu, C. T., Martin, T. L., & Vause, T. (2008). The Assessment of Basic Learning Abilities Test for predicting learning of persons with intellectual disabilities: A review. *Behaviour Modification, 32,* 228–247.

Martin, G. L., & Yu, C. T. (2000). Overview of research on the Assessment of Basic Learning Abilities test. *Journal on Developmental Disabilities, 7,* 10–36.

Martin, G. L., Yu, D., Quinn, G., & Patterson, S. (1983). Measurement and Training of AVC discrimination skills: Independent confirmation and extension. *Rehabilitation Psychology, 28,* 231–237.

Sakko, G., Martin, G. L., Vause, T., & Martin, T. (2004). Visual–visual nonidentity matching assessment: A worthwhile addition to the assessment of basic learning abilities test. *American Journal on Mental Retardation, 109,* 44–52.

Vause, T., Yu, C. T., & Martin, G. L. (2007). The Assessment of Basic Learning Abilities Test for persons with intellectual disability: A valuable clinical tool. *Journal of Applied Research in Intellectual Disabilities, 20,* 483–489.

Yu, C. T., Martin, T., Vause, T., & Martin, G. L. (2015). Kerr-Meyerson Assessment of Basic Learning Abilities-Revised: Recent findings and a conceptual analysis of ordering. *International Journal of Behavior Analysis & Autism Spectrum Disorders, 1,* 55–66.

Received: 30 January 2019
Initial Acceptance: 28 March 2019
Final Acceptance: 10 July 2019

Education and Training in Autism and Developmental Disabilities, 2020, 55(2), 228–242

Effect of Adapted Physical Education and Homework on Gross Motor Development for Young Children with Down Syndrome

Amanda Young, Lisa Silliman-French, and David Nichols
Texas Woman's University

Kathleen Kyzar
Texas Christian University

Abstract: The purpose of this study was to examine the impact of early childhood adapted physical education (APE) and caregiver-implemented homework on the gross motor skill development of 19 young children with Down Syndrome (DS). The results of three groups were compared: (a) APE only, (b) APE plus picture card homework, and (c) APE plus instructional online video homework. The Test of Gross Motor Development-Second Edition (TGMD-2; Ulrich, 2000) was used to determine pretest, posttest, and retention scores following a 12-week intervention. Quantitative data were analyzed using mixed-design factorial ANOVA. All groups made significant gains in object control skill scores; there was no significant difference between groups. Qualitative data analysis (Lincoln & Guba, 1985) of the family questionnaire responses, supported the use of homework for young children with DS.

Educators continue to look for evidenced-based teaching practices to improve the gross motor development of children with disabilities, specifically Down syndrome (DS). With evolving teaching practices and pedagogical knowledge, educators and families can work together to meet the individual educational needs of young children with DS (i.e., adapted physical education [APE], homework). Improving gross motor skills at a young age can promote success in a school environment and help close gaps to reach gross motor milestones.

High quality research, specifically related to APE, comes from a rather small pool in relation to many educational subjects, especially in early intervention. Some researchers started to align APE practices with legislation passed in 1990 (Cowden & Eason, 1991). However, there is a paucity of research to support their proposed practices. Researchers have supported physical activity and play in early childhood and have recommended interventions to increase physical activity for young children (Adamo et al., 2015; Stephens et al., 2013). Researchers have offered evidence to support that young children with DS are capable of participating in physical activities (Barr & Sheilds, 2011; Pitetti et al., 2013; Virji-Babul et al., 2006). Other researchers have offered a solution to early intervention gross motor practices, but focused on the use of related services (e.g., physical therapy, occupational therapy) and exclude instructional services (e.g., APE; Reeder et al., 2011).

Special thanks to KinderFrogs Laboratory School at Texas Christian University. Correspondence concerning this manuscript should be addressed to Amanda Young, Physical and Health Education, Slippery Rock University, 125 Morrow Field House, Slippery Rock, Pennsylvania, 16057. E-mail: amanda.young@sru.edu

Gross Motor Development Interventions for Children with Down Syndrome

Down syndrome, which is associated with a chromosomal anomaly, is the most prevalent cause of intellectual disability (ID; van Gameren-Oosterom et al., 2011). The prevalence of DS is increasing; according to the Centers for Disease Control and Prevention, DS occurs in about 1 in 700 births in the United States (Parker et al., 2010; Zemel et al., 2015). Children with DS have motor and physical limitations related to the genetic anomaly,

which negatively influence gross motor development (Fidler, 2005; Winders, 2013). For instance, individuals with DS may have difficulties participating in physical activity due to decreased kinesthetic awareness, atlantoaxial instability, decreased static and dynamic balance, congenital heart defects, lower muscle tone and strength, joint hyperextensibility, poor postural control, poor vision, poor hearing, and obesity (Palisano et al., 2001; Share & French, 1982; Tudella et al., 2011). Because of these difficulties, children with DS reach developmental milestones significantly later compared to children without DS, starting as early as 3 months of age (Tudella et al., 2011).

Through the process of motor development, children will naturally reach developmental milestones and begin to develop gross motor skills. Motor skills are essential building blocks in preparing individuals for lifelong daily living skills and recreational skills (Gallahue et al., 2012). Motor skills are grouped into three categories: (a) stability, (b) locomotion, and (c) manipulation (Palisano et al., 2001). Stability refers to activities where children must maintain the center of gravity within their base of support (e.g., balance). Locomotion refers to ways in which children move from one place to another (e.g., walk, gallop, leap, jump, skip). Manipulation, often referred to as object control, refers to motor skills where children receive or exert force on an object (e.g., strike, throw, dribble, catch, kick). Through interventions, motor development can be addressed.

Researchers have supported various interventions to increase gross motor development for children with DS conducted in clinical settings (Wu et al., 2007). School-based interventions and home-based interventions are needed to help support gross motor development for young children with DS. One specific intervention that has been suggested for inclusion into an early intervention and early childhood programs related to gross motor skill development is the use of homework (Diaz-Williams et al., 2015). Diaz-Williams et al. (2015) used movement-based homework activities as a way to improve phonological skills for young children with speech impairments. Similarly, investigators of the current study used movement-based homework in addition to APE services to improve the gross motor skills of young children with DS.

Adapted Physical Education

The National Center for Education Statistics (2018) reported that 13% of students in the 2015–2016 school year qualified for special education services under the 2004 Individuals with Disabilities Education Act (IDEA). Part B of IDEA mandates educational services for students with disabilities ages 3 through 21 years, including physical education. Within schools, many students with disabilities qualify for APE, which is physical education that is modified or adapted so that it is appropriate for students with disabilities (National Consortium for Physical Education and Recreation for Individuals with Disabilities, 2008). APE includes the development of physical and motor fitness, fundamental motor skills, and skills in aquatics, dance, individual and group games, and sports (IDEA, 2004). In compliance with federal regulations, physical education must be provided, when appropriate, to students with disabilities. Similar to classroom subjects, APE teachers are encouraged to collaborate with caregivers to continue practicing skills at home.

Homework

For the purposes of this study, physical activity homework is defined as homework activities that involve physical activity, and can focus on a wide variety of knowledge and skills embedded in the physical education/APE curriculum (e.g., fitness, gross motor skills, extracurricular sports, recreation). Articles have been published over the past 35 years about the opinions and possibility of using homework in physical education classes, but few quality research studies have been published related to the use of gross motor homework activities with picture cards for young children (Diaz-Williams et al., 2015; Docheff, 1990; Horvat, 1982). These researchers suggest that picture cards offer additional visual support to caregivers when working with their child in the home environment, as well as, improvments in speech skills. Therefore, the purpose of this investigation was to examine the impact of early childhood adapted physical education

(APE) and parent implemented homework on the gross motor skill development of young children with DS.

Method

Investigators used a mixed methods research design to address the research question. Three groups of participants were compared: (a) APE only, (b) APE plus picture card homework, and (c) APE plus instructional online video homework. Investigators chose to examine four specific object control skills during the interventions: (a) throwing, (b) catching, (c) kicking, and (d) two-handed striking, which are traditionally incorporated into physical education programs.

Setting

This study was conducted within the context of a private, laboratory early intervention preschool for children 18 months to 6 years with DS. The preschool serves primarily young children with DS and other disabilities, but it also includes children without disabilities. Children attended the preschool 5 days/week, for 7.5 hr/day. In addition to instruction in the classroom related to academic and social-emotional development, all students received APE 5 days/week, Speech Therapy (group and individual) 4 days/week, Physical Therapy (PT) consultation 1 day/week, Music Therapy 1 day/week, Oral-Motor Therapy daily, and individual Occupational Therapy (OT) 2 times/month. The preschool serviced 35 children in three classes with a 4:1 student to teacher ratio. In addition, children without disabilities also attended the preschool (n = 8). The children without disabilities serve as peer models and helped to demonstrate age appropriate play skills, facilitate language, and model appropriate behavior. All students at the school were enrolled in one of three classrooms (i.e., Toddler, Pre-K, LEAP), and were divided by developmental level (i.e., cognitive, motor, language), not by age.

Participants and Sampling Procedures

There were primary and secondary participant groups. Primary participants for this study were a convenience sample recruited from the private laboratory early intervention pre-school described in the previous section. The classroom teachers and administrative staff initiated contact with the caregivers of the potential participants by providing a copy of the recruitment email. Participants were 19 students with Down syndrome, and included 10 males and 9 females ranging in age between 2.9 to 6.3 years. The children's primary cargivers were the secondary participants in this study. Caregivers participated by providing an intervention, as well as, answering a caregiver questionnaire that incorporated demographic information and caregiver perceptions regarding the study following the 12-week intervention.

Family participant demographics. One caregiver from each family (17 mothers and 2 fathers) completed a demographic information questionnaire. Most respondents were between the ages of 35 to 44 (n = 15). Multiple ethnicities were represented: Caucasian, African American, Asian, and Hispanic; however, the majority of participants were Caucasian (n = 11). Educational levels for participants and spouses varied greatly; however, the majority of the respondents had a bachelor's degree or higher (n = 15), and most spouses or co-caregivers had a bachelor's degree or higher (n = 14). Most participating families were married (n = 17); two were divorced or separated. Employment of respondents varied; 13 respondents reported that they worked full time, one caregiver worked part time, and five caregivers were not employed. Respondents reported all spouses or co-caregivers worked full time (n = 19). Household income varied between $70,000 and above $150,000; two did not respond. Over half of the families reported household income to be above $100,000 ($n$ = 10).

Families reported residence in urban areas (n = 9), suburban areas (n = 8) and rural areas (n = 2). Family size (living in the same house) varied between three and seven members, and had between one and five siblings (including child with DS); majority of the families had two siblings in their household (n = 10). All caregivers who participated in this study were the biological caregiver to their child with DS. Some families participated in outside services such as speech (n = 5), OT

Figure 1. Students participating in APE class with a kicking focus.

(n = 1), and PT (n = 1); 14 families did not participate in outside services.

Intervention Procedures

First, Institutional Review Board (IRB) approval was obtained. After caregiver consent forms were signed and returned, all primary student participants (N = 19) completed the initial *Test of Gross Motor Development - 2* (*TGMD-2*; Ulrich, 2000) assessment pretest. Second, primary participants were randomly assigned to one of three groups using an iPad attendance application that generated three random groups of six or seven students (A, B, or C).

Group A: APE. Group A (n = 7) served as the control group where students received APE services 5 days/week at the early intervention preschool. Adapted physical education classes were one 30-min group session every school day. All students who attended the preschool received APE services 5 days/week, regardless of participation in this study. Adapted physical education class lessons were structured as follows: (a) stretching/yoga, (b) locomotor warm up, (c) specific skill introduction, (d) specific skill practice and/or, (e) game related to specific skill, and (f) closure (see Figure 1).

Adapted physical education lesson plans were developed and taught by a highly qualified, state and nationally certified APE teacher. Lessons aligned with the *I CAN Physical Education Activities Guide* (Wessel, 1976) to reinforce gross motor development, the *Special Olympics Young Athletes Program Training Guide* (Special Olympics Texas, 2016), and the *Texas Prekindergarten Guidelines* (TEA, 2015). Each week focused on a different gross motor activity or equipment and included the following: scooters, holiday games, noodle games, rings and cones, catching, nutrition, kicking, balloons, throwing, dance, bowling, and striking.

Daily attendance was recorded for each class by a paraeducator. There were a total of 55 APE sessions during the 12-week intervention. Following the 12-week intervention, the APE teacher administered the *TGMD-2* posttest to determine if gross motor skills improved. The APE teacher also conducted a third assessment following a 6-week summer break to measure retention and/or regression.

Group B: APE plus picture card homework. Group B (n = 6) was one of the two treatment groups. Participants in Group B were provided APE services 5 days/week at the early

TABLE 1

Picture Card Homework Intervention Weekly Focus

Weeks	*Lesson Focus*	*Focal Points for Activities*		
1, 6, 11	Throwing	• Throw with assistance • Eyes focused on target • Motion includes hand passing above shoulder	• Almost complete extension for wind-up • Weight transfer • Two-hand throwing slow motion ball • Overhand throwing small ball	• Hip and spine rotation • Follow through • Throwing for distance and accuracy
2,5,9	Kicking	• Kick with assistance • Pendular leg motion • Eyes focused on the ball	• Step forward with non-kicking leg next to the ball • Hip extension and knee flexion • Contact ball with toe or instep • Kicking for distance	• Forward swing of arm opposite of kicking leg • Follow through • Kicking for accuracy • Kickball
3,8,12	Catching	• Catch with arm assistance (i.e., hand over hand) • Eyes focused on the ball with arms ready to receive the ball • Trap or catch ball with hands or arms and chest	• Hands in front of the body, elbows flexed • Extension of the arms in preparation for ball • Bean bag catch	• Contact ball with hands only • Elbows bend as arms absorb the force of the ball • High ball catch • Bounce catch
4,7,10	Striking	• Two-handed sidearm strike with assistance • Bat swings at approximately waist height • Bat swings forward in horizontal plane • Eyes focused on the ball throughout swing	• Dominant hand gripping bat above nondominant hand • Side orientation with bat held behind dominant shoulder • Side striking	• Hip and spine rotation during swing • Weight transfer from back foot to front foot • Follow through • Intermediate tennis/softball striking

intervention preschool, identical to the Control Group A, and caregivers completed assigned picture card homework with their child. Caregivers were asked to practice one skill (i.e., catching, throwing, kicking, striking) 4 days/week for 15 min each day (Diaz-Williams et al., 2015). Each week, the focus was on one specific skill as determined by the APE teacher. Weekly skills were randomized every 4 weeks over the 12-week intervention (see Table 1).

Picture cards. The picture cards were taken from the *I CAN* fundamental skills object control skill curriculum (Wessel, 1976) and the Special Olympics Young Athletes Program (Special Olympics Texas, 2016). Each object control skill had a progression of activities to meet the needs of all participants. The *I CAN* picture cards provided focal points (e.g., step forward on the nonkicking leg with foot landing next to the ball; hip extension and knee flexion), a written description of what to do, cues to say, materials needed (provided by the APE teacher), and a sketch of an individual performing each focal point. The *Special Olympics Young Athletes Activity Guide* provided a description of activities (i.e., kicking for distance), equipment to use, a written description of the activity, a sketch of a child participating in the activity, as well as, helpful observations and tips for caregivers.

Group C: APE plus video homework. Group C (n = 6) was the second treatment group, where students received APE services 5 days/week at the early intervention preschool, identical to the control group, and caregivers completed instructional online video homework with their child. Caregivers were asked to practice one skill (e.g., catching, throwing, kicking) 4 days/week for 15 min each day. Each week the focus was on one specific skill as determined by the APE teacher. Weekly skills were randomized over the 12-week intervention period, identical to the picture card homework group. Homework videos were posted on the APE teacher's personal website and could be accessed at http://amandayoungape.com/index.php/ape-video-homework. At the conclusion of each week, caregivers were instructed to mark progress on an *I CAN* score sheet (identical to the picture card group). At the end of each week, caregivers turned in the log and equipment to the APE teacher; logs served as a compliance check and accountability that caregivers and students completed the homework activities.

Instructional videos. The APE teacher recorded weekly videos to provide caregivers specific instructions on how to practice designated skills. The first video of each skill focus was a detailed explanation of the focal points of each skill, including demonstrations and descriptions of focal points. Following the explanation of the skill was a clip of a child demonstrating a mature pattern of the specific skill in slow motion. Next, the videos included clips of activities for families to practice at home. The activity portion of the video showed the APE teacher and young children demonstrating the sample activities. Throughout the video, specific directions, cues, and reminders were added to assist the caregiver. Videos included modifications and alternative activities for children at varying levels of development.

The focus of the second and third videos was on new activities for families to practice at home during the weekly homework sessions. In addition to the sample activities, a video clip of a mature pattern was included in the weekly video to remind caregivers of the goal of the activities. The videos also had modifications and specific cues and directions that accompanied the new activities. Caregivers were also reminded to refer to the previous videos if they needed reminders of the focal points and skill progression.

Instrumentation and Testing Procedures

Gross motor instrument. The *TGMD-2* is a standardized and highly reliable assessment for children 3 to 10 years old. It is a valid assessment in the areas of content-description validity, criterion-prediction validity, and construct-identification validity. The *TGMD-2* measures 12 gross motor skills in two subtests: (a) locomotor skills and (b) object control skills. The locomotor skill subtests are run, hop, gallop, leap, horizontal jump, and slide. The object control skill subtest are ball skills that include striking a stationary ball, stationary dribble, catch, kick, overhand throw, and underhand roll.

The APE teacher administered the assessment to the participating children individually. The assessments were administered by the APE teacher to ensure the best results and testing environments; the participating children were familiar with their APE teacher and the gym, and were more likely to participate and do their best (Lavay et al., 2015). The assessment was video recorded to allow for unbiased scoring. The actual scoring was completed by three APE students who had successfully passed a graduate course in APE assessment and are currently highly qualified APE teachers in the public-school system. The three students compared assessment scores for inter-rater reliability.

Caregiver questionnaire and qualitative data collection. Caregivers in all groups answered demographic questions to be used for descriptive statistics of participants. Caregivers in the intervention groups completed a questionnaire at the conclusion of the intervention phase regarding their perceptions of the study, personal challenges or successes, and other notable remarks. Further, caregivers were also asked to answer 10 open-ended questions about their participation in this intervention. Sample questions include: (a) "If you were in a homework intervention group, how did you value your time with your child, (b) Is 15 min four times per week a reasonable time to do "homework" with your child, and (c) Did you notice any developmental changes (e.g., improvements in skills) in your child throughout this study?"

In addition to the caregiver questionnaire mentioned previously, additional qualitative data were collected. Throughout the course of this study, the APE teacher used a journal to note field observations of the APE programming, conversations had with caregivers, and questions posed by caregivers. The data were used to capture contextual information during the intervention (e.g., caregiver perceptions/feelings/attitudes, further explanation of activities, student participation). The data were analyzed qualitatively and used to triangulate data from the APE teacher's observations with the *TGMD-2* scores (Lincoln & Guba, 1985).

Research Design and Data Analysis

A mixed method design was used to compare the three groups based on the two parts of the study. The two parts were: (a) Part I: Motor

TABLE 2

Descriptive Statistics of pretest, posttest, and retention test TGMD-2 Object Control Raw Scores

		Control Group (A) (n = 7)	*Picture Card Group (B) (n = 6)*	*Video Group (C) (n = 6)*	*Combined (n = 19)*
Pretest	*M*	12.0	10.8	10.5	11.2
	SD	8.2	5.6	5.9	6.4
Post Test	*M*	12.0	13.0	11.5	12.2
	SD	6.1	3.7	8.3	6.0
Retention	*M*	14.6	14.5	16.7	15.2
Test	SD	8.4	6.9	10.6	8.3

Note. *M* = mean *TGMD-2* score for each group; SD = standard deviation; *n* = sample size.

Skill Development and (b) Part II: Demographic and Qualitative Data.

Part I: Motor skill development. Data from part I of this study were analyzed using a factorial ANOVA (Field, 2009; Portney & Watkins, 2009). The independent variables were the intervention groups (A, B, and C) and time (i.e., pretest, posttest, retention test). The dependent variable was the *TGMD-2* gross motor assessment raw scores. With this data, the APE teacher determined the significant effects and interactions between variables.

Part II: Demographic and qualitative data. Data from Part II of this study were analyzed using descriptive statistics (for demographic data) and qualitative inquiry methods for the observational analyses and data obtained through conversations with caregivers and questionnaire responses. A member check was used after data analysis to ensure accuracy and validity of caregivers' responses (Lincoln & Guba, 1985).

Results

Part I: Gross Motor Skill Development

The APE teacher determined if there were any significant effects and interactions between variables using a 3 × 3 factorial ANOVA with repeated measures on the second factor, and tested at a .05 level of significance. Pretest, posttest, and retention test data were analyzed for descriptive statistics, including mean scores and standard deviations (see Table 2).

The results of the factorial ANOVA indicated a significant effect of time (i.e., pretest, posttest, retention test) on *TGMD-2* object control raw scores ($F = 4.646$; $p = .017$). Although there was no significant effect of intervention type or interaction effect on the object control raw scores, there were significant increases in object control skill scores over time, with the commonality being all participants received APE services. Scores from the pretest were significantly different from retention scores, as determined by a Sidak post hoc test. Visual representation of object control scores over time show a distinct, but not statistically significant, increase in scores between pretest and retention testing (Figure 2). Locomotor scores were not analyzed in this study.

Part II: Demographic and Qualitative Data

Data were analyzed using descriptive statistics for demographic data. Qualitative inquiry methods were used for the observational analyses (Lincoln & Guba, 1985).

Adapted physical education sessions and homework assignments completed. Over the course of the 12-week intervention, there were 55 possible sessions of APE (2,370 min). In addition to APE attendance, caregivers who participated in an intervention group reported homework sessions completed weekly (i.e., four times for 15 min each week). Over the course of the 12-week intervention, there were 48 possible sessions of homework (720 min).

Findings. The findings of Part II of this study are reported according to following themes: (a) presentation of materials, (b) play time with family, (c) staying focused, and

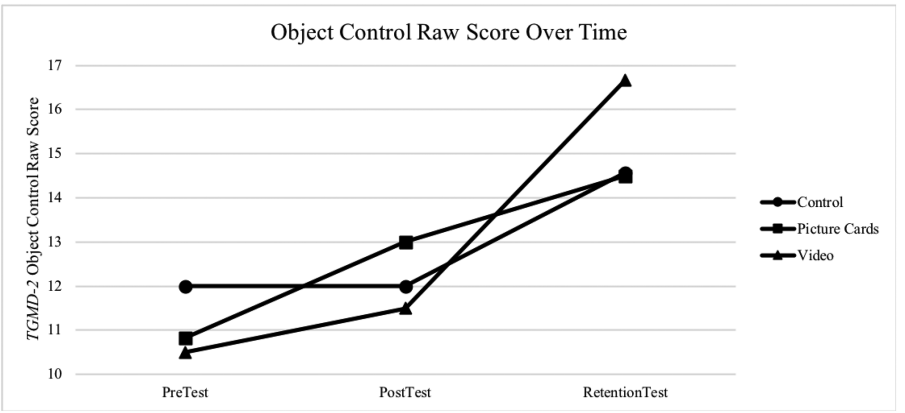


Figure 2. Visual representation of TGMD-2 object control raw score group means over time. Note. Significance occurred between pretest and retention test scores.

(d) developmental progress and gross motor activities. Participants were caregivers who participated in one of the two homework interventions ($n = 12$).

Presentation of materials. Caregivers who participated in homework activities reported likes and dislikes regarding the presentation of homework materials. Caregivers from the picture card group ($n = 6$) reported that they "liked the pictures with step by step instructions" "picture cards were easy to read and refer to quickly" and "the cards were informative on teaching strategies." One caregiver mentioned:

> I love the pictures so you can actually see what position your child should be using to perform the specific activity and then I showed the pictures to [my child] so she could see them too. The picture cards helped me to explain it better too.

Although most caregivers liked the picture cards, one caregiver reported, "picture cards were a bit confusing sometimes." Another caregiver mentioned that they "wished the cards had things to watch for such as [body] alignment." Most responses were positive, with the exception of the two caregivers noting the desire for more information.

Caregivers from the video homework group ($n = 6$) reported that the videos were, "easy to follow," "liked the demonstrations," and "very helpful." One caregiver noted, "there were lots of options and materials for how to complete [the task]." Similarly, another caregiver said, "Videos were very helpful. Every time we had to do the work we were provided with new ideas which were amazingly helpful." In contrast, one caregiver stated, "Sometimes it seemed challenging to take the time to find the video link/log-in and then to watch the video - so that would delay starting the homework. Once at the videos, it was enjoyable to watch them and to see live examples of proper technique." Comparable to the picture card group, most responses were positive.

Play time with family. Numerous caregivers reported the enjoyable time spent with the family completing homework tasks. Caregivers mentioned the following in response to homework activities: "Great family experience," "Well worth the time," "I loved the quality time with [my daughter] and the family," "Fun activity for the entire family," "Helping us improve our quality time together," "I found it to be rewarding and found it fun working with my child," "Very precious and special bonding," "Enjoyable and a good time for the family," "Enjoyed the time spent working on [homework]," "Always the best time when with my kids," "We all enjoyed it," "I loved it and my son too," and "It's a fun way to get the kids moving."

One caregiver noted that her daughter with DS, "loved [homework activities] because it was family fun time and play time for her" and the mother said:

> I loved the time we were able to spend together to those homework assignments and we did it with the grandparents too which became a family fun time session. [My daughter with DS] also loved the extra attention she received from everyone when she accomplished a task.

In addition to questionnaire responses, many caregivers expressed interest in and pleasure about activities directly to the APE teacher frequently during the intervention time period.

Some caregivers mentioned sibling interactions during homework, "[My daughter with DS] has an older sister, I enjoyed watching her try to teach [my daughter with DS] along with us." Similarly, another family noted that homework activities were, "fun for all the kids and me to play with [my son with DS]. We had all three sisters participate made it a family event." Another caregiver reported, "It was a great way to get the family active and to help work on developing motor skills. The kids typically looked forward to Miss [APE teacher's] homework." One caregiver noted, "[We] enjoyed it. My child liked that it was his activity only, similar to his older sibling's daily homework assignments."

Staying focused and prioritizing. One challenging aspect of the homework intervention was the time to complete activities, as well as, keeping the child with DS on task. Behavioral challenges and short attention span for young children with DS are common (Fidler, 2005). Some families mentioned behaviors as challenges to completing homework. One caregiver noted that 15 min of homework was, "feasible, [but] in our case, [our daughter with DS] was the one least cooperative." Similarly, a caregiver said a challenge was, "getting my child to participate." Another caregiver said it was, "sometimes difficult to keep the same exercise exciting for that length of time [15 min]." A few other caregivers also reported difficulties with attention span, "Sometimes maintaining my child's attention for longer than 5 to 10 min was challenging" and "limited attention span by [my daughter with DS]." However, one caregiver mentioned that 15 min of activity was doable, but "any more would be difficult due to limited attention span."

Weather contributed to some difficulties in prioritizing activities. One caregiver stated, "Sometimes it was hard to find the time or the weather made it difficult." Another said, "I think that time frame [15 min] is reasonable - if the weather cooperates . . . on rainy days it was hard and sometimes with other activities we would fall behind." One family mentioned, "We had a lot of rain days during this study. It seemed that [our daughter with DS] worked better outdoors when it was possible." Two other families reported difficulties with weather.

Lack of time and scheduling was another challenge that emerged from the data. Caregivers responded with, "It was more difficult than I expected and there were a couple of weeks that it was impossible because of traveling, sickness, weather, etc." and "It was too much just due to our current situation - being really pregnant then having a new baby, sick kids, etc." In response to homework time requirements, some families responded with personal adjustments they would try, "I think I would have done consecutive sessions and I would have done it more outside of the allotted times" and "Early on in the study I would have treated it like required homework of my older child to prevent those weeks where practice sessions ran into the weekend." Some families make homework activities a priority, and did not feel challenged by time requirements. One family said they would extend homework and combine it with other outside sporting activities and practice. Other caregivers said, "I thought that was the perfect amount of time and I like that it was over 4 days which could span into the weekend if you were busy during the work week" and "the length of time was suitable."

Some families were positive despite challenges, "I thought the overall study was great even though challenging," "it was a somewhat difficult thing to add to our schedule, but worthwhile if it contributes to the body of knowledge on the subject," "sometimes my child did not want to participate but it was fun doing the exercises when he would participate" One caregiver said:

> We had to learn to be inventive and space out the activities so [my daughter with DS] would be able to focus better on her task. She did her best and that was great to see plus she was having fun in the process. Now

> [she] will initiate wanting to throw or kick a ball which is exciting to see.

Overall, most families (n = 8) reported ways to keep their child engaged, and found time to complete homework despite behaviors, weather challenges, and busy schedules.

Developmental progress and gross motor activities. Gross motor skill development observations were noted numerous times in relation to this study. Caregivers shared experiences learning about gross motor skills and helping their child participate in activities. Caregivers said, "It was fun learning," and "It was a good exercise for the caregivers to learn/become more active in play and exercise with their children." One caregiver said, "I did notice that she was more coordinated than I originally thought." Others enjoyed "seeing what my child was capable of doing," and "watching my child progress with the daily practice sessions."

When asked about their child's skill improvement, caregivers reported, "his overall movement became better and I found him becoming more active around the house," "more attentiveness and enjoyed certain activities more than others," "improvement in all areas," and "throwing overhand skills seemed to improve from our first session." One mother said, "I noticed that [my daughter with DS] was able to throw and kick really well. She started to focus a little more when we repeated a task. I saw improvement as the weeks progressed." Another caregiver shared specifically about the focal points of a specific activity, "The only difficulty was trying to get [my daughter with DS] to focus on her follow through and trying to make sure she was using the correct form spelled out on the picture card."

Caregivers also stated their perceptions about their child's engagement in homework activities, "[She] loved it and learned a substantial amount," "It was good to get my son moving around and getting exposed to different ways to play." One caregiver said they, "Enjoyed seeing her excitement" and "I really loved that [she] was enjoying learning." In addition to engagement, caregivers noted the variety and enjoyment of activities, stating that activities were creative and good options for continued practice. A few families mentioned, "We enjoyed baseball and running around the bases, we enjoyed throwing water balloons outside," and:

> I loved learning new ideas/techniques for working on different skills. I never would have thought to tie a wiffle ball on a string to practice hitting or to use crumpled up paper to throw - it was fun to get some insight into the APE activities and to practice together.

In addition, two caregivers noted specific motor and social behavioral improvements. One caregiver said "I enjoyed watching her progress and start to be able to catch and throw and bat. I saw other behaviors improve by adding these interactions." Another noted language development noticing her child speaking new words. Of the 12 families who participated in the homework interventions, 11 reported they would continue to practice gross motor skills at home. One caregiver expressed interest in receiving more activities and picture cards to continue use at home.

In summary, families liked the presentation of materials, they enjoyed being active with their family during homework sessions. At times, focus and attention was difficult, but families generally made homework a priority. Caregivers became aware of gross motor development of their children with DS and enjoyed the variety of homework activities that were presented.

Discussion

The focus of this study was on gross motor development of young children with DS through the use of APE and homework interventions. The results are discussed in the following sections: (a) Object Control Skill Acquisition, (b) Preschool Adapted Physical Education, and (c) Object Control Homework.

Object Control Skill Acquisition

Based on the dynamic systems theory, movements are not predetermined, but influenced by three constraints (i.e., environment, individual, task) which function together (Kugler et al., 1982; Thelen, 2005). The current study

was based on the dynamic systems theory to design intervention programs (serving as task constraints) to investigate gross motor skill acquisition for young children with DS. Newell (1986) categorized the three constraints that interact to generate movements: (a) individual, (b) environment, and (c) task. Specifically, it was noted that the *goal* of the movement was directly related to the movement outcome. Clear movement goals are determined by task analysis of object control skills that are generally used during gross motor play and physical activities (i.e., throwing, catching, kicking, striking). The *TGMD-2* explicitly identifies performance criteria for each skill (Ulrich, 2000). Although clear performance criteria are available, few researchers have tested interventions to address gross motor skill development for young children with DS (Fidler, 2005).

Participants in the current study shared similar individual constraints of age and disability. Children with DS have motor and physical limitations (Fidler, 2005; Winders, 2013); despite limitations, participants in the current study were able to make significant increases in object control skills using the interventions introduced in the current study. Based on the findings, and despite the individual constraints (i.e., decreased kinesthetic awareness, decreased balance, lower muscle tone and strength, poor vision) associated with DS, the participants were able to achieve some or all of the goals of the movements (within the task analysis) related to throwing, kicking, catching, and striking. Over time, significant increases were made for participating students.

Although based on the results, there was a significant increase in skills, it is important to keep in mind that all participating children's *TGMD-2* scores indicated mild to moderate deficits in gross motor development. This relates directly to the findings of other researchers about gross motor development delay for young children with DS (Gémus et al., 2002; Palisano et al., 2001). For instance, Van Duijn et al., (2010) suggested that children with DS acquire motor skills at a slower pace, and reach a "ceiling effect" around 6 years of age. Therefore, it is important to implement evidenced-based gross motor skill interventions at a young age to support the dynamic systems theory.

Preschool Adapted Physical Education

The importance of physical activity for young children with DS is supported by the results of the current study, as well as, a research review by Pitetti et al. (2013). Pitetti et al. reported that children with DS have low cardiovascular and muscular capacity, greater risk for obesity, and do not meet the recommended amounts of daily physical activity. Pitetti et al. also provided recommendations for future research to consider interventions to address the physical activity needs of young children with DS. It is thought that through the support of preschool APE, children with DS may be better equipped to transition to an inclusive kindergarten and elementary school setting by acquiring skills at a faster pace (Van Duijn et al., 2010).

Based on the results of the current study, preschool APE services, as provided during this study, are beneficial to children with DS between the ages of 3 to 6 years. Some researchers have investigated interventions in clinical settings, but no research studies could be located that had been conducted to evaluate interventions in a school-based setting (Winchester et al., 2002; Wu et al., 2007).

Object Control Homework

Gross motor homework activities are supported in opinion literature, but have rarely been used as an intervention strategy in research (Diaz-Williams et al., 2015). In the current study, a similar gross motor homework intervention was used to address gross motor needs in contrast to language needs. With overwhelming support from caregivers, results of the current study indicated that caregivers were willing to participate in homework activities, enjoyed the time spent active with family, and noted developmental progress. Specifically, most families (n = 8) reported child's engagement, and found time to complete homework activities despite challenges.

As reported by participating families, homework completion fidelity was extremely high. The preschool in the current study serves as a university laboratory school; therefore, participating families understand and support the concept of research and are willing to participate in research opportunities. This could

explain the high rate of homework completion, as well as, the possible nature of the homework activities (e.g., active, play in nature).

It is noteworthy that few participating families did not compete all homework sessions; keeping in mind that all families were randomly assigned to a group (i.e., control group, picture card group, video group). The picture card group completed an average of 42.5 homework sessions and the video group completed an average of 46 homework sessions, out of a total of 48 possible sessions. Based on the questionnaire data following the interventions, some families simply could not complete all homework activities. While most families were able to successfully complete homework activities, some families had other obligations that impeded their completion of all sessions.

Fidler (2005) suggested that caregivers should be involved in motor skill interventions to build a strong motor foundation at a young age, which directly supports the homework intervention of the current study. Based on the findings, families enjoyed the opportunity to participate in active homework activities that benefitted their young child with DS. Families noted that the homework sessions were, "rewarding," "fun learning," "great idea," "worthwhile," "love it," and "great even though challenging." Positive reactions from caregivers regarding homework activities support the gross motor development of young children with DS, similar to the recommendations by Pitetti et al. (2013).

Presentation of homework materials using picture cards and videos have been supported in a pilot study using internet-based technologies to train caregivers in working with their child with autism (Meadan et al., 2016). Similarly, caregivers in the current study supported the use of internet-based videos for homework activity instructions. Using videos accessed through the internet provided caregivers a connection to the APE teacher, as well as, concrete, visual, and descriptive examples of how to complete homework activities (van Duijn et al., 2010).

From the positive results of this study, APE teachers are urged to continue addressing skill development at a high frequency for young children with DS. High frequency APE services (i.e., 150 min/week) are a successful intervention, and supported by national physical education guidelines, that can aid in the development of object control skills (National Association for Sport and Physical Education, 2009). Fidler (2005) reported that with multiple repetition of skilled practices, children with DS were able to make significant gains in object control skills.

Implications of the Results and Findings

The results of the current study support preschool APE services for young children with DS. APE teachers should consider providing services to children at the preschool level to help bridge gaps between children with and without DS. If APE services are being provided to children with DS starting at the age of 3 years old, they can gain the skills needed to be successful in a general education physical education environment (Connolly et al., 1993). It is suggested that APE teachers consider multiple teaching styles and activities with appropriate modifications and accommodations to meet the gross motor needs of young children with DS (Mosston & Ashworth, 2008).

Families are encouraged to support gross motor development at a young age through the use of gross motor homework. Gross motor homework at a young age should resemble structured play towards a specific goal. Homework activities outside of school can support motor development for young children with DS. Families are encouraged to find activities their child with DS enjoys and introduce a variety of activities to promote gross motor development which can lead to healthy lifestyles and lifelong physical activity (Young et al., 2016). When families are aware of their child's level of motor development, they can be an essential component in helping to reach motor milestones and increase development (Fidler, 2005).

Limitations

The limitations of this study should be considered in terms of generalizability and evaluation of the results. First, participants represented a convenience sample, therefore limiting the diversity of the participants. Second, caregiver-implemented homework

activities may have been perceived differently for participating families. In addition, the environment and time of day the participants completed their homework with their caregivers could not be controlled. Similarly, multiple family members (e.g., siblings, grandparents) were involved in the homework activities; sibling interaction may have an effect on the homework sessions for the two homework groups (i.e., improper form, distracted practice time). Fourth, potential external factors (e.g., illness, medication, sleep schedules, behavior) may have a negative effect on homework sessions.

This study is also limited by the number of participating families. With a higher number of participants in each group, the sample would have been more statistically robust; however, with a population including individuals with a low incidence disability, sample numbers tend to be low. The early intervention preschool provided a higher number of participants with DS compared to a public preschool setting representing a multitude of disabilities.

Conclusion

Children with DS have motor and physical limitations which significantly influence gross motor development (Fidler, 2005; Winders, 2013). These children may have difficulties participating in physical activity due to decreased kinesthetic awareness, atlantoaxial instability, decreased static and dynamic balance, congenital heart defects, lower muscle tone and strength, joint hyperextensibility, poor postural control, poor vision, poor hearing, and obesity (Palisano et al., 2001; Share & French, 1982; Tudella et al., 2011). Effective gross motor intervention, such as the current study, should be explored to address the individual constraints associated with DS. Concurrently, APE and homework interventions to facilitate gross motor skill improvements should be further explored for young children with DS (Cowden & Eason, 1991).

Within the limitations of this study, it was concluded that APE was an effective intervention used to enhance object control skills for young children with DS. Although not statistically significant, individual data and caregiver opinions support the use of homework as an additional intervention to enhance object control skills for young children with DS. Despite individual constraints related to disabilities, APE and homework tasks can be used across different environments to practice object control skills. The conclusions indicate that APE and homework interventions have a positive effect on gross motor development.

Recommendations for Future Research

Future researchers should replicate this study and examine other object control skills, as well as, locomotor skills. Similar to the current study, a replicated study can be conducted to determine if APE services and homework activities can positively impact additional gross motor skills; and specifically, which evidence-based practices in APE support gross motor skill development. Future researchers should investigate if skills gained in early childhood APE programs and homework activities have a longitudinal effect on children with DS as they enter elementary level of education. Similar to the research conducted by Connolly et al. (1993), researchers should continue to investigate longitudinal studies to support early intervention programming for children with DS.

References

Adamo, E. K., Wu, J., Wolery, M., Hemmeter, M. L., Ledford, J. R., & Barton, E. E. (2015). Using video modeling, prompting, and behavior-specific praise to increase moderate-to-vigorous physical activity for young children with down syndrome. *Journal of Early Intervention, 37,* 270–285. doi. 10.1177/1053815115620211

Barr, M., & Shields, N. (2011). Identifying the barriers and facilitators to participation in physical activity for children with Down syndrome. *Journal of Intellectual Disability Research, 55,* 1020–1033. doi.10.1111/j.1365-2788.2011.01425.x

Connolly, B. H., Morgan, S. B., Russell, F. F., & Fulliton, W. L. (1993). A longitudinal study of children with Down syndrome who experienced early intervention programming. *Physical Therapy, 73,* 170–179.

Cowden, J. E., & Eason, B. L. (1991). Pediatric adapted physical education for infants, toddlers, and preschoolers: meeting IDEA: H and IDEA: B challenges. *Adapted Physical Activity Quarterly, 8,* 263–279. 10.1123/apaq.8.4.263

Diaz-Williams, P., Silliman-French, L., French, R., Nichols, D., & Moorer-Cook, L. (2015). Research

application. Effect of movement-based homework activities on children with a phonological disorder. *Palaestra, 29*(1), 13–16.

Docheff, D. (1990). Homework . . . in physical education? *Strategies, 4*(1), 10–14. 10.1080/08924562.1990.10591754

Fidler, D. J. (2005). The emerging Down syndrome behavioral phenotype in early childhood: Implications for practice. *Infants & Young Children, 18,* 86–103.

Field, A. (2009*). Discovering statistics using SPSS* (3rd ed.). Sage Publications.

Gallahue, D., Ozmun, J., & Goodway, J. (2012). *Understanding motor development: Infants, children, adolescents, adults* (7th ed.). McGraw-Hill.

Gémus, M., Palisano, R., Russell, D., Rosenbaum, P., Walter, S. D., Galuppi, B., & Lane, M. (2002). Using the gross motor function measure to evaluate motor development in children with Down syndrome. *Physical & Occupational Therapy in Pediatrics 21*(2–3), 69–79. 10.1080/J006v21n02_06

Horvat, M. A. (1982). Effect of a home learning program on learning disabled children's balance. *Perceptual and Motor Skills, 55,* 1158. 10.2466/pms.1982.55.3f.1158

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

Kugler, P. N., Kelso, J. S., & Turvey, M. T. (1982). On the control and coordination of naturally developing systems. *The development of movement control and coordination* (pp. 5–78). John Wiley & Sons.

Lavay, B., French, R., & Henderson, H. (2015). *Positive behavior management in physical activity settings,* (3rd ed.). Human Kinetics.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage Publications.

Meadan, H., Snodgrass, M. R., Meyer, L. E., Fisher, K. W., Chung, M. Y., & Halle, J. W. (2016). Internet-based parent-implemented intervention for young children with autism a pilot study. *Journal of Early Intervention, 38*(1), 3–23. doi.10.1177/1053815116630327

Mosston, M., & Ashworth, S. (2008). *Teaching physical education: First online edition.* Retreived from: http://savaria.elte.hu/tmk/spri/hekler/Dokumentumok/SMANTE,%20SMALTE,%20SMDLTE%203303%20A%20testnevel%C3%A9s%20tan%C3%ADt%C3%A1sa/Mosston%20%C3%A9s%20Ashworth%20Teaching%20Physical%20Education.pdf

National Association for Sport and Physical Education. (2009). *National standards and guidelines for physical education teacher education.* American Alliance for Health, Physical Education, Recreation and Dance.

National Center for Educational Statistics. (2018). Children and youth with Disabilities. Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp

National Consortium for Physical Education and Recreation for Individuals with Disabilities. (2008). *APENS: Adapted physical education national standards.* Retrieved from http://www.apens.org/

Newell, K. M. (1986). Constraints on the development of coordination. *Motor Development in Children: Aspects of Coordination and Control, 34,* 341–360.

Palisano, R. J., Walter, S. D., Russell, D. J., Rosenbaum, P. L., Gémus, M., Galuppi, B. E., & Cunningham, L. (2001). Gross motor function of children with Down syndrome: Creation of motor growth curves. *Archives of Physical Medicine and Rehabilitation, 82,* 494–500.

Parker, S. E., Mai, C. T., Canfield, M. A., Rickard, R., Wang, Y., Meyer, R. E. . . . National Birth Defects Prevention Network (2010). Updated national birth prevalence estimates for selected birth defects in the United States, 2004–2006. *Birth Defects Research Part A: Clinical and Molecular Teratology, 88,* 1008–1016. doi.10.1002/bdra.2073

Pitetti, K., Baynard, T., & Agiovlasitis, S. (2013). Children and adolescents with Down syndrome, physical fitness and physical activity. *Journal of Sport and Health Science, 2*(1), 47–57. 10.1016/j.jshs.2012.10.004

Portney, L. G., & Watkins M. P. (2009). *Foundations of clinical research applications to practice* (3rd ed.). Pearson.

Reeder, D. L., Arnold, S. H., Jeffries, L. M., & McEwen, I. R. (2011). The role of occupational therapists and physical therapists in elementary school system early intervening services and response to intervention: A case report. *Physical & Occupational Therapy in Pediatrics, 31*(1), 44–57. 10.3109/01942638.2010.497180

Share, J. B., & French, R. W. (1982). *Motor development of down syndrome children: Birth to six years.* Educational Associates.

Special Olympics Texas. (2016). *Special olympics young athletes program training guide.* Austin, TX: Joseph P. Kennedy Jr. Foundation for the Benefit of Persons with Intellectual Disabilities.

Stephens, T. L., Kinnison, L, Proctor, C., French, R., & Silliman-French, L. (2013). Physical activity: Movement and play in early childhood. *Early Years: An International Journal of Research and Development, 33*(1), 18–23.

Texas Education Agency. (2015). *Texas prekindergarten guidelines.* Retreived from file: ///Users/ajyoung/Downloads/PKG_Final_2015_navigation.pdf

Thelen, E. (2005). Dynamic systems theory and the complexity of change. *Psychoanalytic Dialogues, 15,* 255–283. 10.1080/10481881509348831

Tudella, E., Pereira, K., Basso, R. P., & Savelsbergh,

G. J. (2011). Description of the motor development of 3–12 month old infants with Down syndrome: The influence of the postural body position. *Research in Developmental Disabilities, 32,* 1514–1520. 10.1016/j.ridd.2011.01.046

Ulrich, D. A. (2000). *Test of gross motor development* (2nd ed.). Austin: Pro-Ed.

van Duijn, G., Dijkxhoorn, Y., Scholte, E. M., & Berckelaer-Onnes, V. (2010). The development of adaptive skills in young people with Down syndrome. *Journal of Intellectual Disability Research, 54,* 943–954. doi.10.1111/j.1365-2788.2010.01316.x

van Gameren-Oosterom, H. B., Fekkes, M., Buitendijk, S. E., Mohangoo, A. D., Bruil, J., & Van Wouwe, J. P. (2011). Development, problem behavior, and quality of life in a population based sample of eight-year-old children with Down syndrome. *PLoS One, 6,* e21879.

Virji-Babul, N., Kerns, K., Zhou, E., Kapur, A., & Shiffrar, M. (2006). Perceptual-motor deficits in children with Down syndrome: Implications for intervention. *Down Syndrome Research and Practice, 10,* 74–82. doi.10.3104/reports.308

Wessel, J. A. (1976). *I CAN The field service unit plan in physical education and recreation for the handicapped.* Michigan State University. Hubbard.

Winchester, P., Kendall, K., Peters, H., Sears, N., & Winkley, T. (2002). The effect of therapeutic horseback riding on gross motor function and gait speed in children who are developmentally delayed. *Physical & Occupational Therapy in Pediatrics, 22*(3–4), 37–50. 10.1080/J006v22n03_04

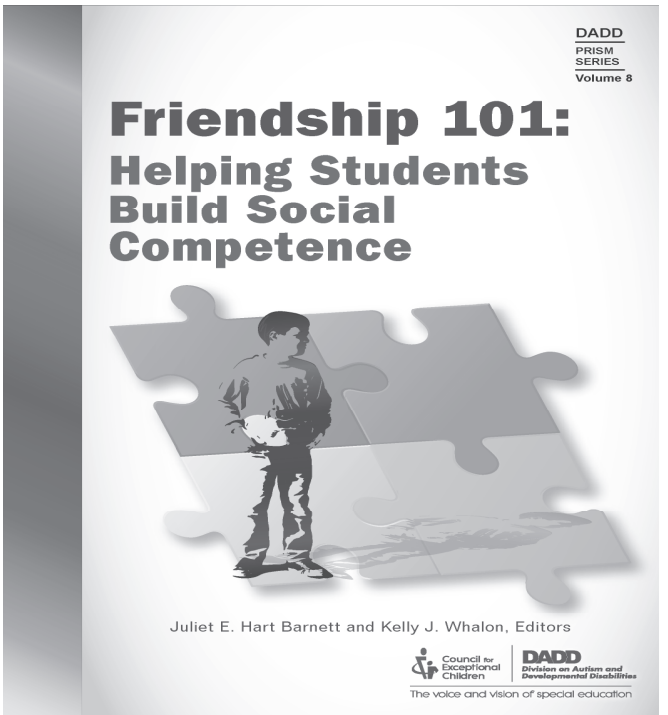
Winders, P. (2013). *Gross motor skills in children with Down syndrome: A guide for parents and professionals* (2nd ed.). Woodbine House.

Wu, J., Looper, J., Ulrich, B. D., Ulrich, D. A., & Angulo-Barroso, R. M. (2007). Exploring effects of different treadmill interventions on walking onset and gait patterns in infants with Down syndrome. *Developmental Medicine & Child Neurology, 49,* 839–945. doi.10.1111/j.1469-8749.2007.00839.x

Young, A. J., Kyzar, K. B., Tolbert, M., & Huckaby, M.F. (2016). After school recreation and families of young children with down syndrome: a qualitative study. *Palaestra, 30*(4), 9–19.

Zemel, B. S., Pipan, M., Stallings, V. A., Hall, W., Schadt, K., Freedman, D. S., & Thorpe, P. (2015). Growth charts for children with Down syndrome in the United States. *Pediatrics, 136,* e1204–e1211. doi.10.1542/peds.2015-1652

Received: 6 March 2019
Initial Acceptance: 23 April 2019
Final Acceptance: 28 June 2019

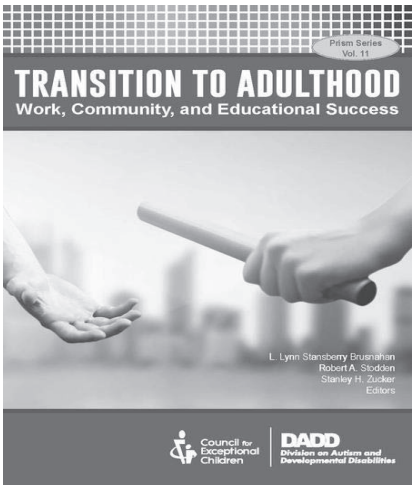


Juliet E. Hart Barnett and Kelly J. Whalon, Editors
2014

The eighth volume of the CEC Division on Autism and Developmental Disabilities' Prism series, *Friendship 101* focuses on building social competence, friendship making, and recreation and leisure skills among students with autism spectrum disorder and other developmental disabilities. Chapters in this evidence-based, user-friendly guide address the needs of students in different developmental periods (from pre-K through young adulthood), providing teachers, parents, faculty and teacher educators with tools and strategies for enhancing the social skill development of these children and youth. Presented through an ecological perspective, together these chapters emphasize building social competence within and across school, home, and community contexts.

Available from CEC Publications
https://www.cec.sped.org/Publications

2900 Crystal Drive, Suite 1000 • Arlington, VA 22202-3557 • 1.888.CEC.SPED (232-7733) • www.cec.sped.org



Edited by L. Lynn Stansberry Brusnahan, Robert A. Stodden, and Stanley H. Zucker. 2018. 149 pages. ISBN 978-0-86586-535-3. #P6296.

The transition from high school to adulthood is challenging for many young people, and often particularly difficult for those with disabilities. *Transition To Adulthood: Work, Community, and Educational Success* provides a blueprint for supporting youth with disabilities in achieving their postsecondary goals in a variety of adult settings – education and training, employment, and the community.

This publication of the Council for Exceptional Children's Division on Autism and Developmental Disabilities is the latest in its successful Prism series (Prism 11) and covers a wide range of topics, from assessing students' interests and abilities to fine-tuning their education plans and goals, ensuring that students with disabilities are included in a variety of settings, and building community relationships to ensure their continuing inclusion. It provides a valuable resource for transition personnel, special and general educators, and special education administrators at the school and state level, as well as adult service professionals.

With eight chapters written by 20 authors, *Transition to Adulthood* covers the breadth of research delineating best practices and proven instructional strategies for ensuring that students with disabilities reach their full potential and achieve their goals.

Visit the CEC online catalog at pubs.cec.sped.org to order this book. The direct link to the book listing is http://pubs.cec.sped.org/p6296/

Education and Training in Autism and Developmental Disabilities

Editorial Policy

Education and Training in Autism and Developmental Disabilities focuses on the education and welfare of persons with autism and developmental disabilities. *ETADD* invites research and expository manuscripts and critical review of the literature. Major emphasis is on identification and assessment, educational programming, characteristics, training of instructional personnel, habilitation, prevention, community understanding and provisions, and legislation.

Each manuscript is evaluated anonymously by three reviewers. Criteria for acceptance include the following: relevance, reader interest, quality, applicability, contribution to the field, and economy and smoothness of expression. The review process requires two to four months.

Viewpoints expressed are those of the authors and do not necessarily conform to positions of the editors or of the officers of the Division.

Submission of Manuscripts

1. Manuscript submission is a representation that the manuscript is the author's own work, has not been published, and is not currently under consideration for publication elsewhere.
2. Manuscripts must be prepared according to the recommendations in the **Publication Manual of the American Psychological Association** (Seventh Edition, 2020).
3. Each manuscript must have a cover sheet giving the names and affiliations of all authors and the address of the principal author.
4. Research studies, including experimental (group and single-subject methodologies), quasi-experimental, surveys, and qualitative designs should be no more than 20–30 typewritten, double-spaced pages, including references, tables, figures, and an abstract.
5. Graphs and figures should be originals or sharp, high quality photographic prints suitable, if necessary, for a 50% reduction in size.
6. Manuscripts along with a transmittal letter and cover sheet that gives the names, affiliations, and address of all authors should be emailed (etadd@asu.edu) to the Editor: Stanley H. Zucker, Mary Lou Fulton Teachers College, Box 871811, Arizona State University, Tempe, AZ 85287-1811.
7. Upon receipt, each manuscript will be screened by the editor. Appropriate manuscripts will then be sent to consulting editors. Principal authors will receive notification of receipt of manuscript.
8. The Editor reserves the right to make minor editorial changes which do not materially affect the meaning of the text.
9. Manuscripts are the property of **ETADD** for a minimum period of six months. All articles accepted for publication are copyrighted in the name of the Division on Autism and Developmental Disabilities.
10. Please describe subjects (or any other references to persons with disabilities) with a people first orientation. Also, use the term "intellectual disability" (singular) to replace any previous term used to describe the population of students with significant limitations in intellectual functioning and adaptive behavior as manifested in the developmental period.

22nd International Conference on Autism,

Intellectual Disability and Developmental Disabilities

"Research-Informed Practice"

January 20 – 22, 2021 in Clearwater Beach, Florida

CLEARWATER
BRIGHT AND BEAUTIFUL · BAY TO BEACH

CEC's Division on Autism and Developmental Disabilities (DADD) welcomes the submission of proposals for innovative and engaging presentations that explicitly link research to practice within the following or related topical areas:

- Evidence-based Practice
- Assessment & Instruction
- Communication & Social Skills
- Self-determination
- Life Skills & Sexuality
- Assistive Technology
- Behavioral Supports
- General Curriculum Access
- Transition Across the Lifespan
- Diversity & Inclusion
- Mental Health
- Coaching & Training

CONTINUING EDUCATION PROVIDER · BEHAVIOR ANALYST CERTIFICATION BOARD · ACE · BACB APPROVED

Council for Exceptional Children | DADD

For further information, please contact Cindy Perras, DADD's Conference Co-ordinator: cindy.perras@gmail.com.