Education and Training in Autism and Developmental Disabilities

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

Volume 54  Number 4  December 2019
Education and Training in Autism and Developmental Disabilities
The Journal of the Division on Autism and Developmental Disabilities,
The Council for Exceptional Children

Editor: Stanley H. Zucker
Arizona State University
Mary Lou Fulton Teachers College

Editorial Assistant: Stephanie McBride-Schreiner
Arizona State University
Mary Lou Fulton Teachers College

Consulting Editors
Martin Agran
Kevin Ayres
Devender Banda
Juliet Hart Barnett
Laura Bassette
Kyle D. Bennett
Emily Bouck
Amanda Boutot
Michael P. Brady
Stacy Carter
Karen Hayden
John McDonnell
Hedda Meadan-Kaplansky
Pam Mims
Wendy Oakes
Jenny Root
Karrie Shogren
Tom E.C. Smith
Scott Sparks
Fred Spooner
Robert Stodden
Keith Storey
William Therrien
Matt Tincani
Jason Travers
Toni Van Laarhoven
Elizabeth West
John Wheeler
Mark Wolery
Leah Wood
Dahun Zhang

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 conveys 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.


Division on Autism and Developmental Disabilities
The Council for Exceptional Children

Board of Directors

Officers

PRESIDENT-ELECT Gina Shadie
PRESIDENT Michael Wehmeyer
VICE PRESIDENT Robert Pennington
SECRETARY Angi Stone-MacDonald
TREASURER Gardner Embarger

Members

Elizabeth Harkins
Nikki Murdick
Cary Trump (Student Representative)

Executive Director
Teresa Doughty
Publications Chair
Michael Wehmeyer
Conference Chair
Emily C. Bouch
Conference Coordinator
Cindy Perras

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.
JENNIFER R. LEDFORD, ERIN E. BARTON, KATHERINE E. SEVERINI, KATHLEEN N. ZIMMERMAN, and ELIZABETH A. POKORSKI

Push Hard and Fast: Teaching College Students with Intellectual Disability to Perform Hands-Only Cardiopulmonary Resuscitation 328
KELLY B. KEARNEY, MICHAEL P. BRADY, CHARLES DUKES, and ANGELICA DOWNEY

Online Module Plus eCoaching: The Effects on Special Education Teachers’ Comprehension Instruction for Students with Significant Intellectual Disability 343
AFTYNNE E. CHEEK, MARCIA L. ROCK, and BREE A. JIMENEZ

Systematic Review of Functional Communication Training in Early Care and Education Settings 358
LI LUO, NICHOLAS A. GAGE, and DEBRA A. PRYKANOWSKI

Increasing Peer-Directed Vocal Manding of Young Children with Autism Spectrum Disorder 377
KAYLA CENTONE, MATT TINCANI, and ABIGAIL DILKS

Comparison of Concrete and App-Based Manipulatives to Teach Subtraction Skills to Elementary Students with Autism 391
LAURA BASSETTE, EMILY BOUCK, JORDAN SHURR, JIYOON PARK, and MCKENZIE CREMEANS

Video Self-Modeling and Functional Behavior Assessment to Modify Aggressive Behaviors in Students with Autism Spectrum Disorder and Intellectual Disabilities 406
KATE M. SADLER

Inclusive Social Studies Content Instruction for Students with Significant Intellectual Disability Using Structured Inquiry-Based Instruction 420
JOANNA RYAN, J. MATT JAMESON, OLIVIA FUDGE COLEMAN, CARRIE EICHELBERGER, JESSICA A. BOWMAN, LYNDSEY AIONO CONRADI, SUSAN S. JOHNSTON, and JOHN MCDONNELL

INDEX—VOLUME 54, 2019 437

Manuscripts Accepted for Future Publication in Education and Training in Autism and Developmental Disabilities 314
Manuscripts Accepted for Future Publication in Education and Training in Autism and Developmental Disabilities

March 2020

Building early numeracy through virtual manipulatives for students with intellectual disability and autism. **Bree Jimenez** and Jessica Besaw, University of Sydney, Sydney School of Education and Social Work, Rm 709, Education Building A35, NSW, 2006, AUSTRALIA.

Investigating the usage of reading curriculum-based measurement (CBM-R) to formatively assess the basic reading skills of students with intellectual disability. **Sara M. Snyder** and Kevin Ayres, James Madison University, College of Education, Memorial Hall 3130E, MSC 6911, Harrisonburg, VA 22807.

Development of the self-determined learning model of instruction coaching model: Implications for research and practice. **Mayumi Hagiwara**, Karrie A. Shogren, Kathleen L. Lane, Sheida K. Raley, and Steven A. Smith, Kansas University, Center on Developmental Disabilities, 3134 Haworth Hall, 1200 Sunnyside Avenue, Lawrence, KS 66045.

Applicability and feasibility of pivotal response treatment with Korean American children with autism. **Sunyoung Kim** and Audrey A. Trainor, University of Illinois at Chicago, Dept. of Special Education, 1040 W. Harrison St., MC147, Chicago, IL 60607.

Perceptions of service providers regarding the agency and capacity to people with intellectual disability to vote. **Martin Agran**, Hannah Ginn, Victoria Estrada-Reynolds, Sandra Root-Elledge, and Eric Moody, University of Wyoming, Special Education Program, Dept. 3374, 1000 E. University Ave., Laramie, WY 82071-2000.

Comparing the effects of echoic prompts and echoic plus picture prompts on establishing intraverbal behavior for children with autism. Sheng Xu, Ruihua Niu, **Gabrielle T. Lee**, Lina Gilic, Hua Feng, and Weiting Shao, 5528 Whitfield Drive, Troy, MI 48098.

Learning fractions with a virtual manipulative based graduated instructional sequence. **Emily C. Bouck**, Courtney Maher, Jiyoum Park, and Abbie Whorley, Michigan State University, 620 Farm Lane, 335 Erickson Hall, East Lansing, MI 48824.


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

Jennifer R. Ledford, Erin E. Barton, and Katherine E. Severini
Vanderbilt University

Kathleen N. Zimmerman
University of Kansas

Elizabeth A. Pokorski
Vanderbilt University

Abstract: Two studies regarding the graphic display of single case data are presented. First, measurement and graphical display of data from studies in highly-ranked journals in special education were assessed. Measurement of desirable behaviors rather than undesirable behaviors was common and most studies used sessions and percentages as x and y-values, respectively. Data presentation (graph ratios and number of data points and sessions per cm) was highly variable; average ratios did not align with previously-published recommendations. In the second study, 50 editorial board members of special education journals were surveyed to determine preferences for graphing ratios. Preferences did not align with recommended graphing practices and varied based on the number of sessions depicted on the graph.

Single case design (SCD) research plays a critical role in establishing evidence-based practices for individuals with developmental disabilities (Horner et al., 2005; Wong et al., 2015). Given its focus on the individual case and repeated measurement that is sensitive to change, SCD research aligns well with the goals of practice in special education and related fields (Ledford, 2018). Thus, assessing the nature and extent of SCD research available to support a specific practice is critical for improving outcomes for individuals with disabilities.

Visual analysis is the primary means by which SCD data are analyzed (Horner et al., 2005; Lane & Gast, 2013; What Works Clearinghouse [WWC], 2017). Visual analysis allows for formative decision-making (Ledford, Lane, & Severini, 2017) and can be used summatively to answer research questions. Specifically, SCD researchers should systematically assess for level, trend, and variability within and across conditions (WWC, 2017). Researchers should also examine consistency and immediacy of behavior change and degree of data overlap across conditions (Ledford et al., 2017; WWC, 2017). To facilitate accurate visual analysis, researchers should use appropriate visual display of their time series data (Morley & Adams, 1991; Tufte, 2001).

Several studies have shown that visual analysis of SCD data is dependent on a number of factors, including variability and magnitude of behavior change (Matyas & Greenwood, 1990), analyst expertise (Ledford & Wolery, 2013), and immediacy of behavior change (Lieberman, Yoder, Reichow, & Wolery, 2010). Factors that may increase the difficulty of SCD data analysis are the use of disparate graphing formats (Kubina, Kostewicz, Brennan, & King, 2017); number of data paths (Hagopian et al., 1997); direction of behavior change (Bartlett, Rapp, & Henrickson, 2011); and use of interval-based recording systems, which have known disadvantageous properties that could influence visual analysis (Ledford, Ayres, Lane, & Lam, 2015; Pustejovsky, 2016). Similarly, studies have shown that data display can considerably influence decision-making, including visual analysis of SCD data (Dart & Radley, 2017; Lauer & Post, 1989). Graphing guidelines include choosing an appropriate format (i.e., a line graph for time series data; Ledford & Gast, 2018) and using consistent height-to-
length ratios (e.g., 2:3, 3:4 or 3:5; Cooper, Heron, & Heward, 2007; Johnston & Pen-
npacker, 1980). However, consistent guidelines
regarding graphing of single case designs are
not available (Ledford & Gast, 2018).

Previous studies regarding graph character-
istics in journals related to behavioral analysis
have concluded that graphs were often pub-
lished absent crucial information (e.g., axis
labels) and with height to length ratios that
were not within previously suggested ranges
and units of measurement (Kubina et al.,
2017). Kubina and colleagues found that only
15% of graphs had ratios falling between their
identified “optimal” range of 63 to 75%. They
also found inconsistencies within and across
studies regarding other graph characteristics
(e.g., length of axes, labeling of axes, consist-
sency of axes values). The extent to which
inconsistencies reflect author or publisher er-
rors, preferences, or lack of published stan-
dards is unclear. Graphing errors and incons-
stancies may be less likely in research with
individuals with disabilities given recently pub-
lished SCD standards by agencies with ties to
special education (e.g., Council for Exception-
tional Children, 2014; WWC, 2017), even
though these standards do not explicitly ad-
dress high-quality visual display of data.

Variability in the graphic display of data, in
conjunction with the aforementioned factors,
may result in inconsistent visual analysis—for
example, a single graph presented at two dif-
ferent ratios might result in differential con-
clusions. Understanding data display and
graphing practices in the field may lead to
increased consistency of use of graphing con-
ventions and subsequent improvements in the
consistency of visual analysis conclusions; this
may be especially critical given the increased
emphasis on synthesis of data across studies.

The purpose of this manuscript is to char-
acterize graphing practices and preferences
for SCD data presentation in special educa-
tion. The purpose of Study #1 was to identify
the state of visual display of data in SCD stud-
ies that meet minimum standards for rigor
(i.e., three potential demonstrations of effect;
WWC, 2017) in journals related to special edu-
cation. Specifically, research questions were:
(1) What design types are widely used in spe-
cial education research? (2) What dependent
variable types are widely measured in special
education research? (3) What height x length
ratios are used? The purpose of Study #2 was
to evaluate the extent to which expert SCD
researchers preferred commonly-cited opti-
mal graphing ratios versus commonly-pub-
lished graphing ratios or other ratios with the
assumption that raters prefer graphs that are
easiest for them to analyze. The research ques-
tions were: (1) What ratio of y-axis to x-axis is
reported as most preferred by SCD experts?
(2) Do preferred ratios vary based on number
of sessions or design type?

STUDY #1

Method

Published SCD studies were located by identi-
fying highly-rated journals in special educa-
tion, identifying which of those journals regu-
larly published SCD studies, identifying all
SCD studies in those journals, and then cod-
ing information for all SCD graphs in each
study. Graphs were located and coded by six
graduate students in special education, all of
whom were working towards eligibility to sit
for the certification exam for applied behav-
ior analysts (Behavior Analysis Certification
Board® [BCBA] exam) and their supervising
professor, an expert in single case design with
a doctoral degree in special education.

First, we identified the top 15 ranked jour-
nals in the “Special Education” category for
Google Scholar and Web of Science. At the
time of the search, eight journals were in-
cluded on both lists (Exceptional Children [EC],
Focus on Autism and other Developmental Disabil-
ities [FADD], Journal of Deaf Studies and Deaf
Education, Journal of Emotional and Behavioral
Disorders, Journal of Learning Disabilities [JLD],
Journal of Positive Behavior Interventions [JPBI],
Journal of Special Education [JSE], Remedial and
Special Education); 7 journals were included in
the Google list only (Education and Training in
Autism and Developmental Disabilities [ETADD],
Education and Treatment of Children [ETC],
Gifted Child Quarterly, International Journal of
Inclusive Education, Learning Disability Quarterly
[LDQ], Teaching Exceptional Children, Topics in
Early Childhood Special Education [TECSE]); and
7 journals were on the Web of Science list only
(American Journal on Intellectual and Developmen-

Following the initial selection of journals relevant to special education, we identified journals which regularly published SCD studies that met the minimum criterion of including at least three potential demonstrations of effect (cf. Council for Exceptional Children, 2014; Ledford & Gast, 2018; WWC, 2017). For example, a multiple baseline (MB) design with two tiers or A-B-A designs were not included, while 3-tiered MB designs (with concurrent baselines and three separate intervention start points) and A-B-A-B designs were included. To determine which of the 22 identified journals regularly published SCD studies, we hand-searched all issues of the most recent complete volume for each journal. Any journal that published two or more eligible SCD studies was included. Twelve journals met this criterion (AJIDD, EC, ETADD, ETC, FADD, JLD, JPBI, JSE, LDQ, RASD, RDD, TECSE).

After identifying eligible journals, a hand-search of each was conducted to identify each eligible article that included a single case design study (using the same criterion listed above) published in the previous five years. Recent publications were identified because graphing practices may be considerably different over time given technological advances, software availability and capabilities, and changing standards in the field. For each study, we separately coded each graph (if applicable). For example, an article including A-B-A-B designs for two participants in two different graphs was coded in two separate rows.

For each identified graph, we coded a variety of elements. First, coders identified the graph design type, including multiple baseline and multiple probe (MB & MP; across contexts, participants, or behaviors); alternating treatments (ATDs); adapted alternating treatments (AATDs); and withdrawal designs (A-B-A-B). Additional variations were coded as “other”, with a descriptive label (e.g., changing criterion design). Coders also recorded the number of data paths and counted the number of data points; if multiple data paths were present, the data points in the longest data path were counted. Intermittent data collection (e.g., generalization sessions) outside of the primary data path were counted as a separate data path if the number of data points was greater than half of the total number of sessions (e.g., if the primary data path had 20 data points and five generalization sessions were conducted intermittently throughout the study, this was not coded as a data path); otherwise, these data points were disregarded.

The intended direction of behavior change was coded based on whether the intervention targeted increasing desirable behaviors (e.g., number of words read) or decreasing undesirable behaviors (e.g., percentage of intervals with aggression). If one graph included multiple data paths intended to change in opposite directions, the coder recorded both. Data collectors recorded the measurement units of the x- and y-axes (e.g., percent, number, rate, duration) based on the axes labels, along with description of measurement in text (if needed). The maximum values of the x- and y-axes were recorded based on the maximum value denoted on each axis. To measure the length of both axes, the graphs were printed at 100% scale on a letter sized sheet of paper and were measured with a ruler (to the closest tenth of a cm). Multiple tiers or panels of a single graph (e.g., MB across participants) were measured separately and averaged. We calculated additional variables using calculations in the Microsoft Excel spreadsheet:

1. Using axis lengths, we derived a ratio by dividing y-axis values by x-axis values.
2. Using the length of the x-axis and the maximum value of the x-axis, we derived a number of sessions per centimeter.

Agreement for coding was conducted for 20% of issues for each journal (i.e., all issues in a randomly selected year), using the previously identified graphs. Agreement was calculated using point-by-point dichotomous assessment for all codes except for measurement of axis lengths, which was calculated by dividing the smaller number by the larger number. Average agreement was 95% (range across journals = 90–99%). Following initial coding some consistent errors were identified for a
single coder, despite high-average levels of agreement. Data from this coder were discarded and the third author re-coded those data. The first author then coded 20% of these articles for the purposes of establishing reliability of the re-coding (95%).

Results
Across journals, 1123 graphs were identified in 267 articles. The number of designs per article ranged from 1 to 54, with a mean of 4 and a median of 2 designs per article. As shown in Table 1, some journals published many single case graphs across 5 years (i.e., ETADD = 277, RASD = 247, RDD = 198). Others published relatively few (i.e., JLD = 5, LDQ = 12).

Research Designs
Table 2 shows graph frequency by design type. There were many ATD graphs (n = 346), MP design graphs (n = 241), and MB design graphs (n = 220). There were fewer A-B-A-B design graphs (n = 133) and AATD graphs (n = 91). Other designs (combined in Table 2) include multitreatment, combination, changing criterion, and simultaneous treatments designs. Because some articles included

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graph Characteristics by Journal</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N (1123)</th>
<th>Data Paths</th>
<th>Data Points</th>
<th>Sessions</th>
<th>Height</th>
<th>Length</th>
<th>Ratio</th>
<th>Sessions per cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJIDD</td>
<td>34</td>
<td>2.3 (1–5)</td>
<td>17.6 (2–95)</td>
<td>21.1 (3–94)</td>
<td>3.3 (2.3–4.6)</td>
<td>8.2 (4.8–15.8)</td>
<td>0.45 (0.25–0.81)</td>
</tr>
<tr>
<td>ETADD</td>
<td>277</td>
<td>1.7 (1–5)</td>
<td>19.2 (3–95)</td>
<td>34.5 (4–190)</td>
<td>4.1 (1.6–12.4)</td>
<td>10.9 (3.7–21.3)</td>
<td>0.39 (0.15–0.99)</td>
</tr>
<tr>
<td>ETC</td>
<td>129</td>
<td>2.0 (1–7)</td>
<td>21.8 (2–185)</td>
<td>32.5 (4–171)</td>
<td>3.7 (1.1–7.9)</td>
<td>8.2 (3.1–12.4)</td>
<td>0.46 (0.11–0.86)</td>
</tr>
<tr>
<td>FADD</td>
<td>68</td>
<td>2.5 (1–4)</td>
<td>14.3 (3–49)</td>
<td>24.9 (6–71)</td>
<td>3.9 (2.1–6.2)</td>
<td>8.6 (4.6–13.5)</td>
<td>0.46 (0.21–0.74)</td>
</tr>
<tr>
<td>JLD</td>
<td>5</td>
<td>1.6 (1–4)</td>
<td>20.4 (17–30)</td>
<td>33.8 (18–90)</td>
<td>3.5 (3.1–4.7)</td>
<td>9.0 (6.6–14.0)</td>
<td>0.42 (0.22–0.48)</td>
</tr>
<tr>
<td>JPBI</td>
<td>86</td>
<td>1.95 (1–4)</td>
<td>21.5 (4–65)</td>
<td>26.6 (6–65)</td>
<td>3.9 (1.7–7.7)</td>
<td>8.9 (4.9–15.8)</td>
<td>0.48 (0.29–0.78)</td>
</tr>
<tr>
<td>JSE</td>
<td>41</td>
<td>1.8 (1–4)</td>
<td>23.7 (8–56)</td>
<td>37.0 (8–94)</td>
<td>4.0 (1.1–6.2)</td>
<td>10.7 (6.7–14.6)</td>
<td>0.42 (0.09–0.93)</td>
</tr>
<tr>
<td>LDQ</td>
<td>12</td>
<td>1.7 (1–3)</td>
<td>21.3 (13–36)</td>
<td>33.7 (15–53)</td>
<td>4.06 (1.4–9.2)</td>
<td>8.20 (5.3–10.4)</td>
<td>0.54 (0.13–1.74)</td>
</tr>
<tr>
<td>RASD</td>
<td>247</td>
<td>1.9 (1–7)</td>
<td>24.6 (2–215)</td>
<td>33.9 (7–215)</td>
<td>3.7 (1.3–7.2)</td>
<td>8.4 (3.5–15)</td>
<td>0.49 (0.17–0.86)</td>
</tr>
<tr>
<td>RDD</td>
<td>198</td>
<td>2.2 (1–7)</td>
<td>26.0 (2–262)</td>
<td>39.7 (4–660)</td>
<td>3.61 (1.5–9.3)</td>
<td>9.17 (1.6–15.3)</td>
<td>0.46 (0.13–2.25)</td>
</tr>
<tr>
<td>TECSE</td>
<td>35</td>
<td>1.4 (1–2)</td>
<td>27.4 (11–46)</td>
<td>42.1 (12–125)</td>
<td>4.1 (2.5–9.4)</td>
<td>10.76 (4.5–15.2)</td>
<td>0.39 (0.17–0.66)</td>
</tr>
<tr>
<td>Average across graphs</td>
<td>2.0 (1–7)</td>
<td>22.1 (2–262)</td>
<td>33.8 (3–660)</td>
<td>3.83 (1.1–12.4)</td>
<td>9.34 (1.6–21.3)</td>
<td>0.45 (0.09–2.25)</td>
<td>4.15 (0.14–28.07)</td>
</tr>
</tbody>
</table>

Note. All cells include mean values with minimum and maximum values in parentheses.

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graph Characteristics by Design Type</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N (1123)</th>
<th>Data Paths</th>
<th>Data Points</th>
<th>Sessions</th>
<th>Height</th>
<th>Length</th>
<th>Ratio</th>
<th>Sessions per cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATD</td>
<td>346</td>
<td>2.91 (1–7)</td>
<td>10.5 (2–83)</td>
<td>20.32 (4–80)</td>
<td>3.83 (1.6–9.3)</td>
<td>7.62 (1.6–14.6)</td>
<td>0.54 (0.18–2.25)</td>
</tr>
<tr>
<td>AATD</td>
<td>91</td>
<td>2.58 (1–3)</td>
<td>19.36 (7–65)</td>
<td>24.86 (11–65)</td>
<td>4.25 (1.5–9.2)</td>
<td>8.84 (4–12.7)</td>
<td>0.54 (0.13–1.74)</td>
</tr>
<tr>
<td>ABAB</td>
<td>153</td>
<td>1.29 (1–5)</td>
<td>30.32 (10–69)</td>
<td>40.89 (10–660)</td>
<td>3.69 (2–7.4)</td>
<td>10.63 (5.1–15.3)</td>
<td>0.37 (0.16–0.71)</td>
</tr>
<tr>
<td>MB</td>
<td>220</td>
<td>1.36 (1–4)</td>
<td>31.36 (8–131)</td>
<td>38.30 (8–131)</td>
<td>3.42 (1.1–6.3)</td>
<td>9.96 (2.1–15.8)</td>
<td>0.36 (0.09–0.95)</td>
</tr>
<tr>
<td>MP</td>
<td>241</td>
<td>1.30 (1–4)</td>
<td>24.82 (5–215)</td>
<td>42.28 (8–215)</td>
<td>3.80 (1.2–12.4)</td>
<td>10.50 (3.7–21.3)</td>
<td>0.38 (0.12–0.99)</td>
</tr>
<tr>
<td>Other</td>
<td>92</td>
<td>2.12 (1–7)</td>
<td>39.14 (6–262)</td>
<td>53.45 (11–260)</td>
<td>4.6 (1.6–9.4)</td>
<td>10.45 (4.6–14.2)</td>
<td>0.46 (0.13–1.00)</td>
</tr>
<tr>
<td>Average across graphs</td>
<td>2.0 (1–7)</td>
<td>22.1 (2–262)</td>
<td>33.8 (3–660)</td>
<td>3.83 (1.1–12.4)</td>
<td>9.34 (1.6–21.3)</td>
<td>0.45 (0.09–2.25)</td>
<td>4.15 (0.4–28.07)</td>
</tr>
</tbody>
</table>

Note. ATD = alternating treatments design; AATD = adapted alternating treatments design; ABAB = Withdrawal design; MB = multiple baseline design; MP = multiple probe design; Other = multitreatment (28), combination (48), changing criterion (13), and simultaneous treatments (3).
multiple graphs, the number of articles including each design was somewhat different, with 119 articles including MB designs, 112 articles including MP designs, 58 articles including ATDs, 42 articles including A-B-A-B designs, 32 articles including "other" designs, and 22 articles including AATDs. MP designs were somewhat balanced between MP designs across behaviors \( (n = 11005 119) \) and across participants \( (n = 11005 108) \), with few MP designs across contexts \( (n = 11005 14) \). MB designs were mostly across participants \( (n = 11005 114) \), with fewer designs across behaviors \( (n = 11005 64) \), contexts \( (n = 11005 8) \), or both participants and behaviors \( (n = 11005 4) \). The majority of the MB designs included three tiers \( (n = 11005 143) \) with 47 including four tiers and 30 including five or more tiers. Similarly, MP designs generally included three tiers \( (n = 200) \), with fewer including four \( (n = 38) \) or more \( (n = 3) \) tiers.

**Dependent Variables**

Most graphs (79%) included data for desirable behaviors, 17% included data for problematic behaviors, and 4% included both types of data. As shown in Table 3, nearly all graphs depicting undesirable behaviors (e.g., challenging behaviors) were ATDs. As expected given their use primarily with non-reversible behaviors, no MP designs or AATDs included measurement of undesirable behaviors, which are typically reversible in nature.

More than half of graphs included a y-value that represented a percentage \( (n = 658) \); other common metrics were count \( (n = 291) \) and rate \( (n = 113) \). Less commonly used were time \( (n = 16) \), rating scale or quality score \( (n = 15) \), cumulative count \( (n = 6) \), ratio \( (n = 2) \), and physiological measures \( (n = 3) \); heart rate, moderate-to-vigorous physical activity). Only 18 graphs had unlabeled y-axes (1.6%). Of the graphs including a dependent variable measured as a percentage, 471 (72%) reported a percentage of opportunities or trials (e.g., percentage correct for academic tasks, percentage of independent responses to social initiations). Most of the remainder (178; 27%) reported a percentage of intervals as an estimate of behavior occurrence. Six graphs reported a percentage derived from a duration measure (percentage of session). Three graphs included multiple data paths with different metrics.

Of the graphs that included use of interval-based systems (IBS; 16%), 132 reported the use of partial interval recording (PIR; 73%), 18 reported the use of momentary time sampling (MTS; 10%), and 21 reported the use of whole interval recording (WIR; 12%). Ten graphs included interval measurement but did not provide the name of the interval system or enough information to determine which was used. Authors used a wide range of interval durations (5-600 s), with an average duration of 22 s and a median duration of 10 s. Graphs using MTS used interval sizes of 10 – 120 s in duration \( (\text{mean} = 41 \text{ s}; \text{median} = 5 \text{ s}) \). Graphs using WIR included interval sizes from 5 – 60 s in duration \( (\text{mean} = 20 \text{ s}; \text{median} = 15 \text{ s}) \). Graphs using PIR included intervals 5 s to 10 min in duration \( (\text{mean} = 20 \text{ s}; \text{median} = 10 \text{ s}) \).

**Data Presentation**

Across research designs, the average number of data paths was 2 \( (\text{range} = 1–7) \). ATDs had an average of 2.91 and AATDs had an average of 2.28. The remaining designs had averages below 2, with the exception of "other" designs (see Table 2). The number of data points per design (or tier, for MB and MP designs) also varied; ATDs included an average of only 10.5

---

**TABLE 3**

<table>
<thead>
<tr>
<th>Design Type</th>
<th>N Desirable</th>
<th>Undesirable</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATD</td>
<td>346</td>
<td>196</td>
<td>150</td>
</tr>
<tr>
<td>AATD</td>
<td>91</td>
<td>91</td>
<td>0</td>
</tr>
<tr>
<td>ABAB</td>
<td>133</td>
<td>106</td>
<td>16</td>
</tr>
<tr>
<td>MB</td>
<td>220</td>
<td>201</td>
<td>7</td>
</tr>
<tr>
<td>MP</td>
<td>241</td>
<td>234</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>92</td>
<td>61</td>
<td>16</td>
</tr>
<tr>
<td>Totals</td>
<td>1123</td>
<td>889</td>
<td>189</td>
</tr>
</tbody>
</table>

*Note.* ATD = alternating treatments design; AATD = adapted alternating treatments design; ABAB = Withdrawal design; MB = multiple baseline design; MP = multiple probe design; Other = multitreatment (28), combination (48), changing criterion (13), and simultaneous treatments (3).
data points, while MB designs included an average of 31.36. Within design types, the number of data points varied widely—the smallest range was 58 data points (AATDs, minimum = 7, maximum = 65). As expected, some graphs included many more “sessions” than data points. This generally occurred in ATDs when each session represented a datum point on a different path and in MP designs, which have missing data by design; both of these design types had nearly twice the number of sessions as data points.

When viewed at 100% (fit to a typical page size), the average graph was 3.83 cm tall and 9.34 cm long, with a ratio of 0.45. The average ratios across journals were relatively similar (all fall within the 0.39–0.54 range) while the average ratios across graph types were slightly more variable. Specifically, A-B-A-B, MB, and MP designs included graphs with smaller ratios (i.e., had longer length given a similar height) and ATDs and AATDs had larger ratios (i.e., had shorter length given similar height). Although average ratios were similar, some journals had extremely wide-ranging ratios. Several journals included ratios that were different by a multiple of 10 (e.g., 0.1 to 1.0), including *JSE, LDQ, and RDD*—this list includes journals from the bottom, middle, and top tertile (third) in terms of total number of SCD graphs. Other journals had much smaller ranges, including *TECSE, JLD,* and *FADD*—these journals are in the bottom or middle tertile in terms of total number of SCD graphs. Only 103 graphs (9.2%) included ratios that were near previously reported ideal ratios, which we identified as those with ratios between 0.6 and 0.7 (i.e., approximating 3/5 to 2/3). As shown in Table 4, most graphs across design types had ratios between 0.25 to 0.5.

Across design types and journals, the average number of sessions per cm was 4.15, with a fairly restricted average across journals (from 2.45 in *AJIDD* to 5.26 in *RASD*), but with a very large range at the individual graph level (range: 0.4–28.07). The number of sessions per cm was fairly consistent across research designs, with the highest numbers for “other” designs (6.74 sessions per cm) and the lowest for A-B-A-B designs (3.74 sessions per cm).

### Table 4

<table>
<thead>
<tr>
<th>Study 1: Graph Ratios by Design Type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ATD</td>
<td>346</td>
<td>0.42</td>
<td>0.49</td>
</tr>
<tr>
<td>AATD</td>
<td>91</td>
<td>0.37</td>
<td>0.52</td>
</tr>
<tr>
<td>ABAB</td>
<td>133</td>
<td>0.24</td>
<td>0.30</td>
</tr>
<tr>
<td>MB</td>
<td>220</td>
<td>0.27</td>
<td>0.35</td>
</tr>
<tr>
<td>MP</td>
<td>241</td>
<td>0.31</td>
<td>0.35</td>
</tr>
<tr>
<td>Other</td>
<td>92</td>
<td>0.33</td>
<td>0.44</td>
</tr>
<tr>
<td>Total</td>
<td>1123</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* ATD = alternating treatments design; AATD = adapted alternating treatments design; ABAB = Withdrawal design; MB = multiple baseline design; MP = multiple probe design; Other = multitreatment (28), combination (48), changing criterion (13), and simultaneous treatments (3).

### Discussion

The measurement and presentation of data in SCD studies in special education is critical for understanding the state of the field. Line graphs are an important tool for evaluating progress, documenting behavior change, and identifying functional relations. Information about the extent to which various SCDs, measurement systems, and graphing conventions are used is relevant for understanding the state of SCD research and evidence-based practice.

### Alignment with Similar Reviews and Recommendations

**Research designs.** Several previous studies have evaluated the extent to which different designs are used in behavioral and special education journals (Hammond & Gast, 2010) and in education and psychology research (Shadish & Sullivan, 2011; Smith, 2012). Although coded somewhat differently across studies, previous reviews have found that time-lagged designs are the most commonly used, followed by withdrawal (A-B-A-B) designs. Hammond and Gast reported the use of MP designs separately, and reported MB designs were used approximately four times as often as MP designs. Interestingly, the current review showed that MB and MP designs were used with approximately the same frequency.
and that ATDs were used most often when measured at the graph level (rather than at the published article level). This suggests that the increasing trend regarding the use of MP designs reported by Hammond and Gast (through 2007) has continued over time.

**Dependent variables.** Several previous studies have evaluated the use of different measurement systems in early childhood special education (Lane & Ledford, 2014), specific areas of special education (functional behavior assessment; Lloyd, Weaver, & Staubitz, 2016), and behavior analytic research (Kubina et al., 2017; Mudford, Locke, & Jeffrey, 2011). Compared with data collected by Kubina and colleagues, a larger proportion of graphs included a dependent variable measuring percentages (59% versus 27%). Compared with data collected by Lane and Ledford and Lloyd and colleagues, a greater number of dependent variables in this broader review were measured via percentage of opportunities or trials and less often measured with interval-based recording systems (16% versus 45% in both previous reviews). Special education researchers might be more likely to conduct studies in typical educational environments (as compared to clinical and research contexts); some measures may be more difficult to conduct in these applied settings. The difference may also suggest that much of SCD research in special education is focused on trial-based behaviors (which tend to be non-reversible, although they are not necessarily so; Ledford & Gast, 2018) rather than free-operant behaviors (which tend to be reversible). Frequent measurement of non-reversible behaviors may be important because much work regarding single case standards and synthesis has focused on designs that are most appropriate for reversible behaviors (e.g., A-B-A-B, ATD, & MB designs; Shadish et al., 2014; WWC, 2017). This suggests that more methodological work is needed for designs associated with non-reversible behaviors. For example, specific visual analysis and internal validity guidelines for MP and AATD designs may be needed. Moreover, because measurement systems could impact functional relation conclusions and effect size estimations, understanding differences in measurement across studies is critical for interpretation.

**Data presentation.** The construction and display of line graphs can vary along critical dimensions, which might impact their usefulness, particularly when synthesizing results across studies. Although graphing conventions have been suggested, with a height to length ratio of 3:5 to 3:4 (i.e., between 0.6 and 0.75; Cooper et al., 2007; Johnston & Pennyacker, 1980), one review showed the percentage of authors approximating this ratio in published graphs in behavior analytic research was only 15% (0.63-0.75 ratio; Kubina et al., 2017). However, Kubina and colleagues did not further report the nature of “errors” from suggested ratios (e.g., range across graphs or journals). The current review suggests the average ratio in published single case research is 0.45, although the range across studies was large. Ratio discrepancies cannot be solely attributed to large numbers of data points given the average number of data points was only 22. Interestingly, many fewer studies in this review (2%) failed to label axes compared to Kubina and colleagues’ review of SCD studies in behavior analytic journals.

**Implications and Future Research Needs**

Outcomes of this review hold several implications for the field. First, measurement of non-reversible behaviors in the context of appropriate designs (i.e., MP, AATD) occurred frequently and as a greater proportion of total SCD studies when compared with other reviews. This is reasonable given that many behaviors expected of students with disabilities are non-reversible in nature (e.g., academic and self-help skills) and may reflect a difference between special education and behavior analytic research. Similarly, most studies measured dependent variables that were intended to increase, suggesting special education research largely focuses on improvement of desirable behaviors rather than reduction of problematic behaviors. In addition, data presentation is highly variable and published ratios (mean of 0.45) do not correspond with previously-suggested standards (0.6 to 0.75). Updated standards may be needed to allow for more accurate cross-study comparison and ease of data interpretation.
Limitations

There are several limitations worth noting. First, we did not code several graphing features, including marker sizes, marker shapes, line weights, tick marks, or graph labels. These features might be important to consider when examining graph characteristics and when developing graphing standards. Second, we intentionally focused on special education journals but might have missed relevant studies. For example, we did not include the Journal of Applied Behavior Analysis, Journal of Behavioral Education, or Remedial and Special Education, which publish a large volume of SCD studies, some of which include individuals with disabilities. Nonetheless, this study provides important data regarding graphing practices in SCD studies.

STUDY #2

Given graphs in special education journals did not follow previously-suggested standards regarding graphing ratios, the purpose of Study #2 was to determine whether experts in SCD preferred ratios consistent with standards (e.g., 0.65, 0.75), publication practices (0.45), or neither. Procedures were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and were approved via the University’s Institutional Review Board. Informed consent was obtained from all participants.

Method

Participants

SCD experts were recruited who served on the editorial board of a journal publishing SCD studies and who had either (a) published 10 SCD studies over the course of their career or (b) published an SCD study in the previous 3 years. First, a graduate student in special education retrieved the names of individuals serving on editorial boards of the 12 journals included in Study #1 (N = 684). Then, the student retrieved publicly-available email addresses for editorial board members. Next, a different graduate student in special education performed a Google Scholar search for each identified email address (n = 630; 54 email addresses not identified) along with the terms “single case” or “single subject.” Individuals with at least three articles identified via this search were recruited (n = 204); in addition, the first two authors hand-searched the list of editorial board members and added 30 participants as potential SCD experts who were not identified via the Google Scholar search. In general, these individuals were not identified because previous SCD research had been conducted and associated with a different email address (e.g., at an institution where the individual was previously employed). A total of 234 recruitment emails were sent; of those, 12 self-identified as ineligible or unavailable. A total of 62 individuals responded and began the survey (28% of 222 potentially eligible recruits). Of these, 50 participants completed the survey in its entirety (81%); their responses were analyzed.

Procedure

Graph generation. Graphs identified during the course of Study #1 were divided into five design types (ATD, AATD, MB, MP, A-B-A-B; no “other” design types were included) and into quartiles based on the number of sessions on the x-axis. One graph was randomly selected from each group (20 total groups, 5 design types × 4 quartiles) using the random function in Microsoft Excel. Then, data from these 20 graphs were extracted using PlotDigitizer (2015), a free program designed for extracting the values of graphed data. Following data extraction, new graphs were generated with Microsoft Excel. All graphs were generated to be 5 cm tall (e.g., y-axis length); to generate graphs with varying ratios of length to height, x-axis lengths were varied from 20 cm (ratio of 0.25) to 6.67 cm (ratio of 0.75). For each original graph, six variations were generated, each with a different y-axis to x-axis ratio (0.25, 0.35, 0.45, 0.55, 0.65, 0.75).

We generated the graphs using a number of rules designed to ensure differences between graphs were relegated to ratios only. For example, given the same ratio, authors might choose to label every session or every other session; this decision might impact analysis of the graph. Rules were: (1) All data provided
by the author in the original graph were included in the generated graphs, including all data paths and any extraneous data points (e.g., generalization, maintenance). (2) All numbers and words were presented in Times New Roman, 12 pt font. (3) The number of unit labels on the y-axis was held constant at 5–8 (e.g., for a range of 0–100%, the values 0, 20, 40, 60, 80, and 100 were used). (4) Data paths were connected with 1.0 pt black lines, and lines were not continued across condition change lines. (5) Filled black circles were used for all data paths; any second data path was generated with unfilled triangles; any third data path was generated with gray-filled squares. (6) Axis tick marks were included for every value for the x-axis (e.g., 10 tick marks = 10 sessions), but were included only for labeled values for the y-axis (e.g., 5–8 tick marks per axis). (7) The number of value labels on the x-axis varied by length, with 1.0–1.6 labels per cm, to ensure similar readability and spacing of labels across ratio presentations. This rule was modified for studies with 33–39 sessions for the 0.25 ratio, because equal spacing and labels every 1.0–1.6 cm was not possible; in this case, a slightly lower rate was used (e.g., labels every 1.75 cm). Similarly, if there were not a sufficient number of sessions for maintaining a label at least every 1.6 cm, every session was labeled (e.g., on a 20 cm axis with 10 sessions, there were only 10 labels, one label every 2 cm).

Survey generation. Following graph development, graphs were saved as PDFs, preserving the original ratios and ensuring consistent quality and presentation across graphs. Each graph set was assigned a letter according to a random function that determined order presentation (e.g., Graph A was presented first). Then, graphs were uploaded using a common web-based survey program. For each graph, five forced-choice two-option items were generated (Graph A with ratio 0.25 versus 0.35, 0.35 versus 0.45, 0.45 versus 0.55, 0.55 versus 0.65, 0.65 versus 0.75). Graph pairs were presented with the question “Which graph presentation do you prefer?” Pairs for each graph were presented in clusters (e.g., Graph A was presented in items 1–5). The first pair to be presented was determined via random selection, and subsequent questions were response-based. For example, if the first item was randomly determined to be 0.45 versus 0.55, and the participant chose 0.55 as their preferred ratio, the next presented pair would be 0.55 versus 0.65. If the participant chose 0.55 again, this was determined to be the most preferred ratio. If the participant chose 0.65, the next item presented would be 0.65 versus 0.75, with the answer to that item corresponding with the participants most preferred ratio. Errors were made in pair presentation or branching logic for two graphs; thus the final number of graph sets to be analyzed was 18. The minimum number of pairs presented per graph was 1 or 2 (based on whether the randomly determined start point was at the extreme of the scale), and the maximum number of pairs presented was 3–5 (again based on the location of the randomly determined start point). The total possible number of graph comparisons presented for all graphs was 32 to 87.

Recruitment. The first author sent a recruitment email to identified editorial board members and a reminder email 10 days later. The email described the purpose of the study and the inclusion criteria, and included a survey link. The survey was available for approximately four weeks after the initial email was sent. Respondents ranged from 28–71 years of age (mean: 48 years). They reported being employed as Assistant Professors (n = 13), Associate Professors (n = 13), or Professors (n = 23), with one participant reporting employment as a practitioner. Most (n = 43) reported Caucasian (Non-Hispanic) ethnicity. Four reported Asian, two reported multiple, and one reported “other” race/ethnicities. Fewer than half (n = 21) reported behavior analysis certifications (BCBA or BCBA-D). Forty-nine reported publishing a single case study in the past three years, and one reported “no” to that question but “yes” regarding whether they had published 10 total studies. Thirty-one participants reported female gender, 17 reported male gender, and 2 did not report gender.

Analysis. Data were downloaded directly from the online program using Microsoft Excel. Descriptive data were calculated and analyzed regarding participant demographics and preferred ratio choices.
Survey Completion

The median time spent on the survey (as reported by the website) was 8 min, 55 seconds. The fastest completion occurred in 5 min; two participants completed in between 45 and 60 min; one participant completed in 100 min; all remaining participants completed in 5 to 33 minutes.

Preferred Ratios

Across graphs and participants, the average preferred ratio was 0.33. Nearly two-thirds (64%) of total responses indicated a preferred ratio at the minimum value (0.25; see Table 5). Participants rarely chose 0.75 (3%), 0.65 (5%), or 0.55 (5%) as the most preferred ratio and sometimes chose 0.45 (10%) or 0.35 (13%). Individually, respondents’ average preferred ratio ranged from 0.28 to 0.50, with 58% of respondents having a preferred average between 0.25 and 0.35, 32% having an average between 0.35 and 0.45, and 10% having an average between 0.45 and 0.55. Each person, with one exception (98%), chose 0.25 as their most preferred ratio for at least one graph and 36% of respondents chose 0.75 as the most preferred ratio for at least one graph.

Variability in Preferred Ratios

Average preferred ratios varied based on starting pair; that is, respondents tended to choose smaller ratios when presented with those first. However, even when the largest ratio pairs were initially presented (e.g., 0.55 versus 0.65 and 0.65 versus 0.75), the average preferred ratio remained low (0.40 and 0.36, respectively). Likewise, ratios varied by design type,
with lower preferred ratios for MB and MP designs (0.30 for both), followed by A-B-A-B designs (0.32), AATDs (0.36), and ATDs (0.42). As shown in Figure 1, much of the variability can be explained by the number of data points; a large portion of respondents consistently chose the minimum value (0.25) as preferred when the graph included 20–40 sessions. When the number of sessions was fewer than 20, the proportion of respondents choosing the minimum value was lower and the average preferred ratio was higher. Three outliers appear for graphs with the largest number of data points; these graphs were all randomly assigned to an initial pair of 0.65 versus 0.75, which may partially explain the discrepancy.

**Discussion**

Notable findings from Study #2 are that (1) self-reported expert preference is even more divergent from previous guidelines than publication practices; (2) regardless of design, experts prefer low ratios (i.e., longer lengths [x-axes] with shorter heights [y-axes]); and (3) experts generally preferred lower ratios (longer lengths) when there was a larger number of sessions. The three outliers, all of which had many data points, also happened to have been graphs with a randomly selected initial ratio comparison of 0.65 versus 0.75. Thus, additional research is needed to confirm the finding that preferred graphing ratios vary according to the number of sessions. In future
research, initial graph presentation should be randomly selected separately for each participant, rather than overall random selection to allow for more fine-grained analysis of this potential relation.

Limitations

The participation rate was somewhat low; the recruitment email included information regarding inclusion criteria (i.e., SCD expertise), so it may be likely that a number of the originally-recruited individuals were not eligible and did not participate due to ineligibility, artificially deflating participation rates. In addition, inaccurate pairs or branching logic was used for two graphs; thus data from those graphs were discarded. Finally, because many responses indicated a preference for the lowest ratio (0.25), it is likely that even lower ratios may be preferred by some individuals. We chose not to use lower ratios given that 0.25 was considerably lower than published recommendations, but additional research is needed including a wider range of ratios. Additional research is also needed regarding consistency in graphing details (e.g., marker size, line weight, and use of arrows to label data series); these may be important for comparing graphs across studies and authors.

Proposed Guidelines

Given results from these studies, other scientific data, and historical standards, we propose several guidelines for the field.

1. Given visual analysis is the recommended method of data analysis for SCD research by which functional relations are identified, quality features of the graphical display of data should be incorporated into SCD standards and visual analysis protocols. Consistent, clear, and accurate graphical data displays might improve the consistency with which visual analysis is conducted and reduce Type I errors.

2. In addition to development of evidence-based standards for graphical display, that the use of publication guidelines is needed to improve the consistency of publication practices across journals. All included journals had minimal guidelines for figures; most commonly, journals only had guidelines related to resolution, although RASD and RDD also included guidelines related to font type as well as uniformity of font and sizing.

3. Although specific ratios have been proposed, it may be important to consider graph size relative to the number of data points. We suggest that ratio sizes between 0.25 and 0.55 may be appropriate for most graphs, with smaller ratios for graphs with a larger number of data points.

4. Sessions are the most commonly used x-axis value. Others have indicated the x-axis should represent calendar time (i.e., calendar days; Kubina et al., 2017) rather than sessions, arguing that the term “sessions” is not time-based. However, sessions are time-based, in at least an ordinal sense (Session 1 occurred before Session 2 in time, and both occurred prior to Session 3). We recommend continued use of sessions at the current time; empirical data are needed that evaluate the use of sessions versus calendar days and whether either adversely impacts visual analysis. X-axis scales should be consistent across graphs within a study and should accurately represent the number of sessions. Authors should take care to accurately describe the unit(s) of time for each session, the time between sessions (e.g., twice per day with two hours in between, once per day, every other day), and the total duration of the study (e.g., study sessions occurred across three calendar months).

Conclusion

SCD research and the subsequent use of visual analysis play a critical role in establishing evidence-based practices in special education (Gast & Ledford, 2018; WWC, 2017). In SCD research the graph is a critical tool used to make both formative and summative decisions. However, our results point to a divergence between graphing standards, published graphs, and expert preferences, which might impact reliability across visual analysts, decisions about functional relations, and analysis of results across multiple studies. As SCD research continues to advance to meet current research standards, standards related to the
graphical display of data also should progress and evolve.

References


Received: 25 July 2018
Initial Acceptance: 20 September 2018
Final Acceptance: 25 October 2018
Push Hard and Fast: Teaching College Students with Intellectual Disability to Perform Hands-Only Cardiopulmonary Resuscitation

Kelly B. Kearney, Michael P. Brady, Charles Dukes, and Angelica Downey
Florida Atlantic University

Abstract: Many adults with intellectual disability (ID) do not learn the skills needed to maintain safety of people within their communities. In a pair of studies, a total task presentation with a least-to-most error correction procedure was used to teach four college students with ID how to safely remove personal protective equipment (PPE) (gloves), and how to administer hands-only cardiopulmonary resuscitation (CPR). A multiple probe design was used to evaluate the effects of both studies. Results for the first study showed that students’ accuracy with PPE removal increased after the intervention, and maintained after the intervention was removed. During the second study, student accuracy of hands-only CPR increased, and the CPR skills also generalized to a novel setting. Social validity data for both studies indicated that students believed they learned and were able to demonstrate the skills, however staff impressions were more variable. Implications and limitations are discussed.

Sudden cardiac arrest is one of the leading causes of death in the United States (American Heart Association, 2018). According to the American Heart Association, almost 90% of people who experience cardiac arrest outside of a hospital die, however, immediate cardiopulmonary resuscitation (CPR) can double or even triple a person’s chance of survival. CPR is the act of placing the heel of the hand over the sternum and compressing the area at a rate of 100–120 beats per minute, with a compression depth of 5–6 centimeters (Khan & Weston, 2014). The American Safety and Health Institute (ASHI), the American Heart Association (AHA), and others have been training citizens to perform Basic Life Support skills (BLS) (skills that provide care to injured people until they can receive full medical attention) such as first aid and CPR, since the 1970s. In 2011, AHA published a statement calling for CPR training in elementary and secondary schools for all students to be mandatory, however this has not yet become a universal practice (Cave et al., 2011). As of 2018, only 38 states have any type of legislation regarding the inclusion of CPR instruction into the high school curriculum (AHA, 2018). According to Cave et al. (2011), the standards vary greatly between the states that have some sort of CPR component in their curriculum. Some states require students simply to “recognize” the steps of CPR, while a few states actually require CPR certification for high school graduation.

CPR is just one of many safety skills that should be taught across the lifespan to people with and without disabilities. Safety skills, which involve knowing both how to prevent accidents and how to respond if faced with an emergency, at home or in the community, are necessary to function as an independent member of society (Collins, Wolery, & Gast, 1992; Dixon, Bergstrom, Smith, & Tarbox, 2010; Ozkan, 2013). Safety skills include a variety of skills and routines needed to maintain physical wellbeing, such as crossing a street, contacting emergency services, and responding to a house fire (Collins et al., 1992; Dixon et al., 2010; Garcia, Dukes, Brady, Scott, & Wilson, 2016). Hands-only CPR is a lifesaving method advocated by the AHA because

This research was funded in part by a grant from the TAFT Foundation to the Florida Atlantic University Academy for Community Inclusion. Correspondence concerning this article should be addressed to Kelly B. Kearney, Department of Exceptional Student Education, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431. E-mail: KBROWN65@fau.edu
it can be performed by someone with minimal first aid training. Hands-only CPR is effective in circulating blood flow, and is advocated to be used by responders who might not otherwise participate in resuscitation due to fear of transfer of bacteria or germs. Chest compressions are effective for the first few minutes someone is in cardiac arrest, pushing oxygenated blood through the body to keep vital organs alive until someone with advanced skills and treatment options arrives (Benjamin et al., 2017).

Although CPR has been recommended by AHA to be universally taught in schools, this skill set frequently is omitted from high school safety curricula. Often students with ID leave high school unprepared to respond to emergencies in the community. Safety skill instruction for this population of students tends to be neglected in favor of more commonly utilized daily living skills, such as communication and social skills (Dixon et al., 2010). Yet personal emergencies, particularly injuries, require attention as people with disabilities are at an increased risk to sustain such injuries (White, McPherson, Lennox, & Ware, 2018). As more students with ID enter postsecondary education programs, it may be beneficial to include BLS skills as part of the college and university curriculum, since employers frequently require employees and interns to be familiar with BLS skills. These programs can empower students with life-saving community safety skills, increasing their community participation in integrated workplaces and public spaces.

A number of teaching methodologies have been studied in an effort to explore various BLS skills, such as first aid, with a wide age range of children and adults with intellectual disability. Teaching procedures for first aid skills include modeling (Spooner, Stem, & Test, 1989), interactive story-telling (Marchand-Martella, Martella, & Marchand, 1991), backward chaining (Gast, Winterling, Wolery, & Farmer, 1992), peer teaching and tutoring formats (Marchand-Martella et al., 1992), small group instruction (Timko & Sainato, 1999), peer and self-video modeling (Ergenekon, 2012; Ozkan, 2013), and literacy-based behavior interventions (Kearney, Brady, Hall, & Honsberger, 2017). Many of these teaching approaches include some type of chaining procedure, but none have employed total task presentation, a method frequently demonstrated as effective in teaching basic self-care skills (Browder & Spooner, 2011; Snell & Brown, 2011).

The BLS studies that targeted individuals with ID included individuals ranging from 3- to 20-years-old. The empirical research on CPR, an important basic life support skill, has largely been limited to typically developing schoolchildren. In 2013, Plant and Taylor reviewed literature to determine effective CPR training methods for adolescents, and again total task presentation was missing from the findings. In 48 articles that were identified and analyzed, Plant and Taylor determined that with behavioral practice students as young as 4 years old could determine whether a person was unconscious and/or breathing erratically, place the person in the recovery position, open the person’s airway, and call emergency services. Plant and Taylor recommended using hands-on training for the physical tasks and repetition of the skills effectively to teach CPR skills to students.

The purpose of this research was to explore the efficacy of an instructional procedure, specifically total task presentation combined with an error correction procedure consisting of least-to-most prompts, to teach two safety routines to college students with ID. This is the first known empirical study using any instructional procedure to teach adults with ID to perform CPR. Two separate studies were conducted. Experiment 1 addressed a common but critical component of most first aid interventions: mastery of the personal protective equipment (PPE) routine of removing and disposing of gloves. Two research questions were addressed:

1. Will the instructional procedure increase the acquisition of a PPE routine (gloves removal) by college students with ID?
2. If the college students acquire the PPE routine, will they maintain their skills after the intervention is removed?

Experiment 2 extended the investigation by incorporating the PPE skills into a life-saving routine of hands-only CPR. The research questions for Experiment 2 were:

1. Will the same instructional procedure in-
increase the students’ ability to perform a CPR routine?
2. If students acquire the CPR routine, will they maintain their skills after the procedure is removed?
3. If students acquire the CPR routine, will the CPR routine generalize to a different setting?

To evaluate the social validity of across both studies, researchers assessed student and staff views regarding students’ ability to perform the skills, the likelihood of their use of the skills in the future, and the appropriateness of the intervention.

**Experiment 1: Method**

**Participants**

Four adults with intellectual disability (ages 21–45) who attended a post-secondary education program for students with intellectual and developmental disabilities at a state university in the southeastern United States participated in both studies. All students were diagnosed with an intellectual disability as their primary eligibility category on their most recent psychological evaluations. All students graduated from high school with a special diploma or certificate of completion, and all had four semesters of participation in college courses leading to a post-secondary certificate in supported employment. Eligibility for the study required: (a) being in good academic standing, (b) regular attendance, and (c) prior completion of a 2-credit university class, *Health and Fitness for Life*, offered for students enrolled in the program. The participants were recommended by their instructors to participate based on student interest in learning the skill and future need based on employment interests. Although all students were age 21 or older, only Gary was his own legal guardian and could provide his own consent to participate. The other students provided assent to participate, and their parents provided consent prior to beginning the study. University institutional review board approval was obtained.

**TABLE 1**

<table>
<thead>
<tr>
<th>Student (Age)</th>
<th>Assessment Data</th>
<th>Disability</th>
<th>Reading Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary (21)</td>
<td>KBIT: 52</td>
<td>Intellectual Disability</td>
<td>Recognition grade 1st; Comprehension grade: 1st</td>
</tr>
<tr>
<td></td>
<td>Performance = 59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support = 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gary (45)</td>
<td>KBIT: 50</td>
<td>Intellectual Disability</td>
<td>Recognition grade 3rd; Comprehension grade: 3rd</td>
</tr>
<tr>
<td></td>
<td>Performance = 70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support = 63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Julio (26)</td>
<td>WAIS: 43</td>
<td>Intellectual Disability</td>
<td>Recognition grade 2nd; Comprehension grade: 1st</td>
</tr>
<tr>
<td></td>
<td>Performance = 79</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support = 77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tim (23)</td>
<td>WAIS: 40</td>
<td>Intellectual Disability</td>
<td>Recognition grade 1st; Comprehension grade: 1st</td>
</tr>
<tr>
<td></td>
<td>Performance = 68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support = 70</td>
<td>(Down Syndrome)</td>
<td>Speech and Language Impairment</td>
</tr>
</tbody>
</table>

*Note. WAIS = Full Scale IQ score from the Wechsler Adult Intelligence Scale-IV (Wechsler, 2008). KBIT-2 = Composite IQ from the Kaufman Brief Intelligence Test-Second Edition (Kaufman & Kaufman, 2004). JOBS: OSD = Job Observation and Behavior Scale: Opportunity for Self Determination (Brady, Rosenberg, & Frain, 2006). For adults in competitive, supported, and sheltered employment, the mean composite score for Quality of Performance = 77.99; the mean composite score for Type of Support required to achieve that performance = 76.17.*
prior to conducting the study. See Table 1 for a summary of student demographics.

All of the investigators held graduate degrees in special education. However, germane to the focus of this intervention, two of the investigators were active and nationally certified first aid and CPR instructors or instructor trainers for the American Safety and Health Institute. One of these investigators delivered all of the intervention sessions.

**Setting**

Experiment 1 took place in the faculty suite or the adjacent hallway in the College of Education building on the campus that the students attended. The site had immediate access to a sink and trash cans in a kitchenette that was readily available to faculty and staff, and both the faculty suite and hallway afforded ample space for the students, materials, and observers. All baseline, intervention, and maintenance observations took place in this setting. No faculty or staff other than the investigators were present in any of the immediate settings.

**Task and Materials**

Prior to Experiment 1 a probe of a universal precaution procedure (safely removing soiled gloves) was administered to each student. This procedure included wearing personal protective equipment, and handling soiled materials appropriately (Occupational Safety and Health Administration, 2018), procedures necessary to avoid infection when handling fluids and tissues from another person. Instruction on glove removal is a routine first aid skill because it is necessary to minimize potential blood and fluid splatter, prevent contamination from soiled gloves, and limit inadvertent exposure to bloodborne pathogens (ASHI, 2016). All materials required to remove and dispose of gloves, and to clean up after disposing of them, were present in the adjacent kitchenette. Gloves used in this study were purchased from a local pharmacy. All gloves used were large, vinyl gloves.

**Behavioral Measure, Data Collection, and Inter-observer Agreement**

A task analysis comprised of eight steps was created by the two ASHI certified investigators for safe removal and disposal of soiled gloves (see Table 2). The task analysis incorporated a consensus protocol based on the 2016 first aid guidelines (ASHI, 2016). To collect data, each student was individually observed while removing the gloves. A data collector rated steps as (a) correct and independent, (b) correct and required a prompt from the researcher, or (c) incorrect or no attempt made. A prompt from the researcher was defined as telling or re-telling the student what to do, modeling the behavior, or a physical assist with hand over hand prompts. The least intrusive prompt (telling the student what to do) was used first, and the next level of prompt was used only if the student still did not perform the step correctly. The most intrusive prompt (the physical assist) was never actually used during the study. Only steps that were correct and independent were graphed, and all decisions involving condition changes were made based on those data. Data were collected by two observers using data sheets created by the researchers. Observers were experienced special education teachers enrolled in a graduate program in special education. Both were trained to record both tasks during live practice sessions with staff who per-

### Table 2

<table>
<thead>
<tr>
<th>Safe Glove Removal and Disposal Task Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Steps for Safe Glove Removal and Disposal</strong></td>
</tr>
<tr>
<td>1. Pinch and hold outside of glove near wrist.</td>
</tr>
<tr>
<td>2. Pull down, away from wrist, turning glove inside-out.</td>
</tr>
<tr>
<td>3. Pull glove away until it is removed from hand, holding inside-out glove with gloved hand.</td>
</tr>
<tr>
<td>4. Use ungloved hand to slide fingers under the wrist of the remaining glove. Do not touch the outer surface of glove.</td>
</tr>
<tr>
<td>5. Pull down, away from wrist, turning glove inside-out.</td>
</tr>
<tr>
<td>6. Continue to pull glove down and over the inside-out glove being held in gloved hand.</td>
</tr>
<tr>
<td>7. Throw gloves in trash.</td>
</tr>
<tr>
<td>8. Wash hands with soap and water.</td>
</tr>
</tbody>
</table>

---

Push Hard and Fast / 331
formed the skills prior to initiating the studies. Observations began only when the observers reached 100% agreement on each behavior code.

The observers also collected inter-observer agreement (IOA) data. During the study, observers stood behind and to the side of the students so they could see and hear directions to the students, and observe the students’ performance, but could not observe one another. Observer agreement was determined by counting the number of steps scored the same on glove removal task analysis, dividing that by the total number of steps on the task analysis, and then multiplying by 100. Inter-observer agreement data were collected during 42% of Mary’s sessions, 38% of Gary’s sessions, 42% of Julio’s sessions, and 42% of Tim’s sessions. Agreement across all students and sessions was 100%.

**Experimental Procedures**

**Baseline.** Baseline included a no-instruction condition during which an investigator began each session by distributing gloves to the student, asking the student to put the gloves on, and then requesting the student to perform the glove removal routine to the best of his or her ability. No additional assistance was provided to the student. During baseline observations, the relevant materials for each experiment (gloves, trash can) were available but no direction was provided on their use. No additional assistance was provided, and baseline sessions ended when students indicated verbally or physically that they were done.

**Intervention.** The intervention was a total task presentation, with error corrections provided using a least-to-most prompting hierarchy (Browder & Spooner, 2011; Snell & Brown, 2011). Total task presentation involves training the student in each step of the chain for every learning trial (Browder & Spooner, 2011; Miltenberger, 2016). Intervention sessions were delivered individually for each student. When delivering the intervention, an investigator implemented a total task presentation, teaching each step of each task analysis from the beginning. The intervention was delivered by first modeling the behavior, then completing a side-by-side walk through (i.e., simultaneously performing the steps with the student), then asking the student to complete the whole task independently. After the first intervention session, each subsequent observation session began by asking students to perform the skill prior to the total task presentation, and data were collected on that performance. This resulted in a minimum of a 24-hour delay for the observation after the previous intervention, thus avoiding the performance data on that day from being influenced by immediate practice effects. Each session started with the student putting gloves on, and then the researcher asking the student to, “Please remove your gloves.”

If a student made an error on any step, the observer marked the step as incorrect, and the investigator immediately prompted the student using a least-to-most prompting hierarchy (Browder & Spooner, 2011; Snell & Brown, 2011). The prompt hierarchy for corrections included: (a) telling the student how the step should be performed, (b) modeling the correct way to perform the step, or (c) providing physical assistance in completing the step. To implement the correction procedure, the investigator pointed out the error, then delivered a prompt, and asked the student to perform that step again. If the student made a subsequent error on the same step, the next level prompt was delivered. Thus, only the least amount of assistance was delivered to assist students to correct errors on any steps on their task analyses.

**Follow-up.** Follow-up probes were conducted following removal of the intervention to assess whether the students would continue to perform the PPE removal routine learned during the intervention. Follow-up sessions were conducted between 37 and 55 days after the last intervention session.

**Experimental Design and Analysis**

A multiple probe design across participants was used to determine the impact of a total task presentation on the PPE routine of glove removal. Multiple probes were used during baseline to prevent lengthy, erroneous practice of the skills prior to intervention (Kennedy, 2005). Follow-up observations were held after the intervention was removed to determine potential maintenance of the learning.

Data were analyzed using traditional visual
inspection procedures, as well as calculating measures of central tendency for the dependent measures during baseline, intervention, and follow-up. Condition changes were made based on the level and trends of individual data points, with decisions to move from baseline to intervention based on multiple data points showing low and stable performance. Decisions to move from intervention to follow-up were based on stable performance, with a minimum of three stable data points demonstrating one error or less.

In a post-hoc analysis, the Percent of Non-Overlapping Data (PND) was calculated, a common effect size estimate for single subject design studies, to supplement the visual inspection of graphed data (Ledford, Wolery, & Gast, 2014; Scruggs & Mastropieri, 2013). PND was established separately for baseline-to-intervention, and from baseline-to-follow-up conditions. Effectiveness ratings were based on standards established by Scruggs and Mastropieri (2013), who defined interventions as (a) highly effective when 90–100% of data do not overlap with baseline, (b) moderately effective when 70–90% of data do not overlap with baseline, (c) minimally effective when 50–70% of data do not overlap, and (d) ineffective when 50% or less of data fall below baseline. A Tau-U coefficient was also calculated to establish a true effect size estimate (Parker, Vannest, Davis, & Sauber, 2011). For this analysis, the Tau-U web-based calculator was used (Vannest, Parker, Gonen, & Adiguzel, 2016). These calculations provided an omnibus effect size based only on the baseline to intervention effect, and did not include the baseline to follow-up or intervention to follow-up differences.

Experiment 1: Results

The effects of the intervention on glove removal are shown in Figure 1. A summary of performance changes across conditions is shown in Table 3.

Baseline

During baseline, data remained low and stable, with no student accurately completing more than 25% of the steps in the glove removal task analysis. Gary consistently completed two of the eight steps correctly. Tim started baseline by completing two steps accurately, but then independently switched his technique and completed zero steps accurately. Julio vacillated between performing zero steps and two steps accurately. Mary did not complete any steps accurately during baseline.

Intervention

After the intervention was implemented, all students showed drastic increases in skill accuracy, with the first intervention session ranging from 88% to 100% accuracy. Mary reached consistent 100% accuracy after two sessions. Gary reached 100% accuracy during the first intervention session, dropped to 88% for the following two sessions, then increased and remained steady at 100% for the rest of the sessions. Julio increased to 100% on the first intervention session, then decreased to 88% for one session, then increased and remained at 100% for the rest of the sessions. Tim jumped to 100% accuracy and remained there after the first intervention session.

Follow-up

Upon the removal of the intervention, Mary and Gary’s skill accuracy remained at 100%. Julio and Tim’s skill accuracy declined during follow-up, decreasing to 88% for both follow-up sessions.

Effect Size Differences across Conditions

A post-hoc analysis of the findings using Percent of Non-Overlapping Data indicated that the intervention was highly effective. The effect size for all students between baseline and intervention was 100%. Between baseline and follow-up, student PND was 100%. The post-hoc analysis using the Tau-U calculator showed an aggregate effect size of 1.00. This suggests a strong effect size across the students (Parker et al., 2011), but it should be noted that this analysis did not include data reported during the follow-up observations.
Figure 1. Percentage of correct independent steps of task analysis for safe glove removal.
Experiment 2: Method

Approximately six weeks following the end of Experiment 1, the same participants returned to campus for the fall semester. Experiment 2 was initiated approximately two weeks into the fall semester.

Participants

The same four students who participated in Experiment 1 also served as participants in Experiment 2. The same investigator who delivered the intervention in Experiment 1 also delivered the intervention in Experiment 2.

Setting

For Experiment 2, the intervention sessions took place in a conference room in the College of Education. The conference room was around the corner from the setting of Experiment 1, approximately 150 feet away. The conference room was roughly $25 \times 12$ feet and contained a large rectangular table with several chairs. Generalization probes in Experiment 2 took place in a teaching assistant office in the same building. The office was roughly $12 \times 12$ feet and contained three separate work spaces with computers and chairs. No faculty or staff other than the investigators were present in the conference room or the generalization site during the study.

Tasks and Materials

The task selected for Experiment 2 was hands-only CPR, a method of CPR advocated for citizen first-responders with minimal first aid training (AHA, 2018). Although hands-only CPR does not include mouth-to-mouth rescue breathing, it does require the use of the PPE routine (safe glove use and removal) targeted in Experiment 1. Because all students demonstrated mastery of that safety routine during that investigation, they were eligible to participate in Experiment 2. All materials required to perform hands-only CPR were present in the conference room and office. All gloves used were large, vinyl gloves, identical to the gloves from Experiment 1, and were purchased from a local pharmacy. Two CPR manikins (Laerdal - Little Anne CPR Training Push Hard and Fast / 335

### TABLE 3

#### Experiment 1 Gloves: Performance Changes Across Conditions and Effect Sizes

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline</th>
<th>Intervention $^1$</th>
<th>Follow-Up $^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>Mean</td>
<td>0%</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0%</td>
<td>(88–100%)</td>
</tr>
<tr>
<td></td>
<td>PND</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
<tr>
<td>Gary</td>
<td>Mean</td>
<td>25%</td>
<td>96.6%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>(25%)</td>
<td>(88–100%)</td>
</tr>
<tr>
<td></td>
<td>PND</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
<tr>
<td>Julio</td>
<td>Mean</td>
<td>10%</td>
<td>97.6%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>(0–25%)</td>
<td>(88–100%)</td>
</tr>
<tr>
<td></td>
<td>PND</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
<tr>
<td>Tim</td>
<td>Mean</td>
<td>4%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>(0–25%)</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>PND</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
</tbody>
</table>

**Note.** Percentage of Non-overlapping Data (PND) contrasts compare Baseline to Intervention $^1$, and Baseline to Follow-up $^2$. 

---

*Push Hard and Fast / 335*
Manikin, model 120–01050) were present in the room during all sessions. One manikin was on the floor during baseline, generalization, and follow-up probes. Two manikins were on the floor during intervention sessions. The gloves were within eye sight of the door, on a table. Manikins were cleaned after each session per ASHI instructor standards (ASHI, 2016).

Behavioral Measure, Data Collection, Inter-observer Agreement, and Fidelity

An 11-step task analysis based on the 2016 consensus protocol guidelines (ASHI, 2016) was created for the delivery of hands-only CPR (see Table 4) by the ASHI certified investigators. As in Experiment 1, data were collected for each student individually during the CPR routine. The task analysis required that steps be completed in sequence. Data were collected for the generalization measure using the same task analysis. As in Experiment 1, data collectors rated steps as (a) correct and independent, (b) correct and required a prompt from the researcher, or (c) incorrect or no attempt made, and any prompts from researchers were defined as telling or re-telling the student what to do, modeling, or providing physical assists with hand over hand prompts. Only steps that were correct and independent were graphed, and all decisions involving condition changes were made based on those data.

<table>
<thead>
<tr>
<th>Steps for Hands-Only CPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assess the scene</td>
</tr>
<tr>
<td>2. Attempt to wake victim</td>
</tr>
<tr>
<td>3. Tell someone to call 911</td>
</tr>
<tr>
<td>4. Put on gloves</td>
</tr>
<tr>
<td>5. Begin chest compressions (30)</td>
</tr>
<tr>
<td>6. Open airway</td>
</tr>
<tr>
<td>7. Repeat chest compressions (30)</td>
</tr>
<tr>
<td>8. Open airway (person starts breathing and wakes)</td>
</tr>
<tr>
<td>9. Place person in sitting up position</td>
</tr>
<tr>
<td>10. Remove gloves</td>
</tr>
<tr>
<td>11. Wash hands with soap and water</td>
</tr>
</tbody>
</table>

The same observers from Experiment 1 collected data and conducted inter-observer agreement observations in Experiment 2, and the same data collection and analysis procedures were used. Observers collected agreement data during 47% of Mary’s sessions, 59% of Gary’s sessions, 57% of Julio’s sessions, and 50% of Tim’s sessions. Observer agreement across all students and sessions was 99%. In addition, inter-observer agreement data were collected in the generalization setting during observations that correspond to intervention, and follow-up conditions. Agreement during these generalization assessments averaged 97.8% and ranged from 91% to 100%.

Fidelity of treatment was assessed during Experiment 2 by an observer who watched the investigator deliver the intervention. To assess treatment fidelity, the observer recorded whether the investigator delivered instruction for each step of the CPR task analysis. The observer placed a check mark next to the corresponding step as it was observed, based on the occurrence or non-occurrence of the researcher behaviors. The observer was trained to collect fidelity data by watching an in vivo session of the researcher performing the behavior chain (without any students present), then discussing the scores with the researcher. No disagreements were noted during the training session.

Across all four students, the investigator delivered a total of 34 intervention sessions. A fidelity assessment was conducted on the delivery of the intervention to students during 17 of these intervention sessions, however these assessments were not distributed evenly across the students. For example, no fidelity assessments were conducted during Mary’s seven intervention sessions. For Tim, six of his 12 intervention sessions (50%) received a fidelity assessment. For Gary, six of eight sessions (75%) received a fidelity assessment, and for Julio five of seven sessions (71%) received a fidelity assessment. The fidelity assessment for each session for these students was 100%.

Experimental Procedures

Baseline. Baseline included a no-instruction condition during which an investigator began each session by escorting each student...
to an area with a manikin positioned on the floor, and saying “I don’t think he’s breathing, can you do CPR?” No other instruction was provided. During baseline observations, the relevant materials (gloves, trash can, manikins) were available but no direction was provided on their use. No additional assistance was provided, and baseline sessions ended when students indicated verbally or physically that they were done.

**Intervention.** The same total task presentation with an error correction of least-to-most prompts was employed for the CPR training sessions in Experiment 2. This involved presentation of the total task, simultaneous practice of the task, and having the student perform the task independently. Data were collected as soon as the student arrived in the conference room so that there was a 24-hour delay after the most recent intervention session, thus preventing immediate practice effects. As in baseline, each intervention session started with “Look, I don’t think he’s breathing, can you do CPR?”

The manikin was positioned on the floor in the same area as it was during baseline. Gloves were also in a similar place. During the side-by-side walkthrough of the task, a second manikin was added. The researcher and student completed the chain facing each other, each using separate manikins. The researcher modeled “push hard and fast” and counting out loud to 30 during compressions, but counting aloud was not required of the student to complete the step accurately during data collection.

**Follow-up and generalization.** Follow-up probes were conducted following removal of the intervention to assess whether the students would continue to perform the safety routine learned during the intervention. Follow-up sessions were conducted between 7 and 74 days after the last intervention session. No additional instruction or prompting was provided during the follow-up observations. The same prompt of “I don’t think he’s breathing, can you do CPR?” was used during both the follow-up and generalization sessions. During observations in the generalization setting, gloves were placed on a desk close to the door and the manikin was on the floor in the center of the room. No intervention took place during generalization sessions.

**Experimental Design and Analysis**

To assess the intervention effects of the CPR routine, a multiple probe design across participants was used. Follow-up observations were conducted after the intervention was removed to determine potential maintenance of the skill. Generalization probes were conducted in each condition to establish whether students could perform the CPR routine spontaneously in another setting.

As in the first experiment, traditional visual inspection procedures were used to examine graphed data and decisions to move from intervention to follow-up were based on stable performance, with a minimum of three stable data points demonstrating less than one error. A percentage of non-overlapping data (PND) was calculated in a post-hoc analysis separately for baseline-to-intervention, and from baseline-to-follow-up conditions using the same effectiveness ratings applied in Experiment 1. Finally, a Tau-U coefficient was calculated using the Tau-U web-based calculator for an omnibus effect size based on the baseline to intervention effect, and excluding the baseline to follow-up or intervention to follow-up differences.

**Experiment 2 Results**

The effects of the intervention on hands-only CPR are shown in Figure 2. A summary of performance changes across conditions is shown in Table 5.

**Baseline**

During baseline, data remained low and stable, with no student accurately completing more than 27% of the steps in the hands-only CPR task analysis. Gary started baseline with zero percent accuracy, however he noticed the gloves present during the third probe and correctly put them on, removed them, and then washed his hands, increasing his accuracy to 27%. Gary returned to 0% accuracy during the fourth baseline probe. Mary, Julio, and Tim did not complete any steps accurately during baseline.
After the intervention, all students showed increases in skill accuracy, with the first intervention session ranging from 64% to 73% accuracy. Julio reached 100% after four sessions. Mary and Gary reached consistent 100% accuracy.

**Figure 2.** Percentage of correct independent steps of task analysis for hands-only CPR.
racy after five sessions. After four sessions, Tim reached 82% accuracy. Tim vacillated between 82% and 100% for seven sessions before remaining at 100% accuracy.

**Follow-up**

Upon the removal of the intervention, Gary and Tim’s skill accuracy remained at 100%. Mary and Julio’s skill accuracy decreased somewhat during follow-up. Mary and Julio remained at 91% during all follow-up probes, consistently forgetting to say “call 911” during all follow-up probes.

**Effect Size Differences across Conditions**

A post-hoc analysis of the findings using PND indicated that the intervention was highly effective. The effect size for all students between baseline and intervention was 100%. Between baseline and follow-up each student’s PND was also 100%. The Tau-U post-hoc analysis showed an aggregate effect size of 1.00. This indicates a strong effect size across all students (Parker et al., 2011), but, as in Experiment 1, it should be noted that this analysis did not include data reported during the follow-up observations.

**Social Validity**

The investigators created a social validity measure to assess the perspectives of the students and university staff on the students’ ability to perform both routines taught in the two experiments. The survey contained five items for university staff, and five items for students. Staff and students were provided with the survey, asked to circle the answer that best represented their perspectives, and then return the survey to an investigator. Response options for each item on the staff instrument were based on a 4-choice Likert scale (Strongly Agree – Agree – Disagree - Strongly Disagree) with a Not Sure option. Response options for each item on the student instrument were based on a 4-choice scale (Absolutely – Kind of – Not really – No way). Data analysis included calculating means for each item. The social validity assessment was conducted two weeks after the last follow-up observation in Experiment 2.

All four students voluntarily participated in the social validity assessment. In addition, four staff members participated in the social valid-

---

**TABLE 5**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Baseline</th>
<th>Intervention</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>Mean 0%</td>
<td>94%</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td>Range (0%)</td>
<td>(73–100%)</td>
<td>(91%)</td>
</tr>
<tr>
<td></td>
<td>Effect size</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
<tr>
<td>Gary</td>
<td>Mean 7%</td>
<td>93%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Range (0–27%)</td>
<td>(64–100%)</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>Effect size</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
<tr>
<td>Julio</td>
<td>Mean 0%</td>
<td>92%</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>Range (0%)</td>
<td>(64–100%)</td>
<td>(91–100%)</td>
</tr>
<tr>
<td></td>
<td>Effect size</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
<tr>
<td>Tim</td>
<td>Mean 0%</td>
<td>88%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Range (0%)</td>
<td>(64–100%)</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>Effect size</td>
<td>–</td>
<td>100% Highly Effective</td>
</tr>
</tbody>
</table>

*Note. Effect size reported as percentage of non-overlapping data (PND) by comparing Baseline to Intervention, and Baseline to Follow-up.*
ity assessment. The staff members worked in student services as academic advisors and administrative assistants and were regularly available in the College of Education building. Although they were knowledgeable of the postsecondary program for students with ID, they did not work directly with the students.

Social Validity Results

The strongest social validity findings with the staff indicated that CPR and glove removal are important skills for the students to learn (M = 3.5 out of 4.0) and that the students were willing to perform CPR (M = 3.7). Three staff members indicated they were not sure if students knew how to perform CPR and remove gloves properly, while one staff member reported that students did know how to perform CPR and remove gloves appropriately (M = 3.0). Two staff members reported they strongly agreed the procedures used to teach the skills were acceptable (M = 4.0), while the other two staff members were not sure. All staff members reported they were not sure if the students wanted to learn other safety skills and routines. Although social validity findings with staff were variable, students were less ambiguous in their perceptions. Students reported that they (a) knew how to perform CPR and remove gloves appropriately (M = 4.0), (b) they were willing to perform CPR (M = 4.0), (c) they wanted to learn other first aid skills and routines (M = 4.0), (d) they thought CPR was important to learn (M = 4.0), and (e) they liked the way they learned CPR (M = 4.0).

Discussion

The purpose of this study was to determine whether college students with ID would learn a BLS skill (glove removal with hands-only CPR). The procedure used, a total task presentation instructional procedure with a least-to-most prompting error correction, had not previously been used to teach BLS skills. All students who received the instructional procedure mastered both skills, and their learning gains maintained for several weeks after the procedure was removed.

To date, this is the first empirical study using any instructional procedure to teach adults with ID to perform CPR. All students mastered both safety routines with 100% accuracy, requiring between four and seven intervention sessions in Experiment 1, and seven and 12 intervention sessions in Experiment 2. These results are promising; although researchers have reported that many caregivers assume older students with disabilities “are incapable of learning” safety skills (Agran & Krupp, 2010, p. 304), this study demonstrates otherwise.

Although students maintained high accuracy levels of the PPE procedures during follow-up after Experiment 1, overall the students did not display this skill during baseline probes in Experiment 2. We hypothesize this may be due to the break between semesters; the baseline probes for Experiment 2 occurred between two and three months after the conclusion of Experiment 1. Perhaps this large gap in time was detrimental to the maintenance of the skill. Also, the researchers did not point out what materials were in the room before providing the prompt to start the chain; perhaps the students did not notice the gloves were present and this affected their performance during baseline.

When examining the data for Experiment 2, it should be noted that most students consistently had a difficult time with two particular steps in the task analysis: (a) attempting to wake the victim, and (b) telling someone to call 911. We can only speculate as to why these two steps were most difficult to master. Perhaps these skills were unfamiliar to the students. Prior to the study all students were aware of CPR from television or movies. Such media presentations frequently portray actors administering CPR with chest compressions, but often the viewer doesn’t see the actor try to wake the person or tell a bystander to call 911. It is possible that such incidental means contributed to the difficulty or ease of learning certain steps.

Continuing to examine the results from Experiment 2, Julio and Mary both mastered all steps in task analysis, and were able to maintain the accurate pace of 100–120 beats per minute when performing hands-only CPR. Tim and Gary’s pacing when administering chest compressions varied, and their latency between hearing the discriminative stimulus “I don’t think he’s breathing, can you do CPR?”
and actually beginning chest compressions was much greater than Julio and Mary. Although accurately completing all steps in the task analysis, it is possible Tim and Gary may not respond as quickly in an emergency situation as Julio and Mary.

Hands-only CPR delivered by a layperson drastically increases the victim’s chances of survival (Cave et al., 2011). Ultimately, during an episode of cardiac arrest, some action is better than none, although the performance of high-quality chest compressions will have a significant impact on the recovery of the victim (Cave et al., 2011). This, and other safety skill training that people with ID in the community receive is likely to increase their independence and help them transition from sheltered to competitive work placements (Mechling, 2008). Although regulations vary by state, many places of employment encourage or even require CPR knowledge or certification.

Although encouraging, the findings should be considered with the following limitations. First, limited fidelity data were collected, with no fidelity data collected during Experiment 1, and with only some sessions during Experiment 2. Future research should include more frequent assessments of treatment fidelity to ensure that the intervention is administered as intended. Second, the generalization probes were in the same building as the intervention sessions. Future researchers might consider probing for generalization in very different settings, such as in a different building, in a parking lot, or in other community settings. The limited generalization probes do not allow researchers to determine the applicability of these skills to situations outside of the university. Third, gloves were provided to students. Outside of study conditions, students will not typically have access to medical gloves before encountering an emergency situation.

This study extends previous research demonstrating that adolescents with intellectual and other developmental disabilities can master basic life support skills (Kearney et al., 2017) by showing the impact of an intervention on adults attending a college program for students with ID. Adults living in natural, integrated communities need to shop for themselves, prepare their own food, and get around the community, but safety skills are often overlooked by educators and caregivers (Dixon et al., 2010). Safety skill routines are important for adults with ID, particularly adults attending college. Individuals who are able to contribute to the health and well-being of those around them are valued community members. Responding to an emergency situation, such as cardiac arrest, by administering hands-only CPR is a desirable community safety skill. Equipping college students with ID with the safety skills and routines needed to respond competently to these types of emergencies will better enable them for typical current and future workplace, home, and community settings.

References


Dixon, D. R., Bergstrom, R., Smith, M. N., & Tar-


Received: 20 September 2018
Initial Acceptance: 14 November 2018
Final Acceptance: 15 December 2018
Abstract: Successful text comprehension is a complex task, requiring a variety of cognitive skills and processes that support students in making inferences, integrating information across sentences in text, monitoring understanding, and gaining knowledge about text structure. For students with significant intellectual disability (SID), text comprehension is a difficult skill to acquire because it is often associated with overall language ability. In this study, we used a multiple baseline across participants design to investigate the effects of an online module plus eCoaching on special education teachers’ (SETs) use of the CROWD (Completion, Recall, Open-ended, WH, Distancing; Whitehurst et al., 2004) in the CAR (Comment, Ask, Respond; Cole et al., 2002) text comprehension strategy during literacy instruction, and the impact on students with SID listening comprehension and engagement. Results confirmed the online module plus eCoaching was effective in increasing special education teachers’ knowledge and use of the CROWD in the CAR comprehension strategy, improving students’ listening comprehension, and maintaining high levels of student engagement.

Historically, literacy instruction for students with significant intellectual disability (SID) has been grounded in two approaches, developmental and functional. When adopting a developmental approach, the special education teacher (SET) emphasizes a subset of foundational literacy skills, which serve as prerequisite skills that students with SID must master prior to engaging in more sophisticated instruction (Kliwer & Biklen, 2001). By comparison, when taking a functional approach, the SET focuses on providing instruction in skills needed for navigation within the school and community, namely sight word instruction (Copeland & Keefe, 2007). Both approaches have been criticized for a lack of explicit, contextualized instruction in text comprehension (Copeland & Keefe, 2007).

For nearly a decade, the need for SETs to provide students with SID explicit, contextualized instruction in text comprehension has been proffered by Browder and colleagues (2009). More recently, Spooner and Browder (2015) reiterated the importance of providing text comprehension to students with SID by stating that “so much of academic learning and adult functioning relies on comprehension of text” (p. 30). Researchers defined text comprehension as a student’s demonstration of understanding of text-based content regardless of their ability to read independently (Knight & Sartini, 2015). Text comprehension is a complex academic task that requires students to employ a variety of cognitive skills and processes, such as making inferences, integrating information across sentences, monitoring understanding, and gaining knowledge about text structure (Cain, Oakhill, & Bryant, 2004).

Moreover, because text comprehension is often associated with overall language ability (Snyder, Knight, Ayers, Mims, & Sartini, 2017), it is a difficult skill, not only for many...
students with SID to learn, but also for SETs to teach. Students with SID often demonstrate expressive and receptive language needs (Browder & Spooner, 2011; Orlando & Ruppar, 2016), as well as other unique challenges that impede active participation in literacy lessons. Some may experience physical challenges while holding books, or when viewing standard print (Erickson, Hanser, Hatch, & Sanders, 2009), and others may require augmentative or alternative communication (AAC) devices to access print (Orlando & Ruppar, 2016). For example, AAC systems can be programmed with vocabulary, to assist students in questioning, predicting, and commenting on text; therefore, increasing meaningful participation, during literacy instruction. Because students with SID have difficulty gaining, maintaining, and generalizing knowledge and skills (Browder & Spooner, 2011), many remain at an early literacy level, relegated to learning skills that lead to reading, throughout their academic careers (Browder et al., 2009).

Model for Effective Literacy Instruction

In 2009, Browder and colleagues proposed a model for effective literacy instruction designed specifically to meet the unique needs of students with SID. A foundational tenant of this model is that literacy outcomes for students with SID should lead to increased access to literature and increased independence as a reader. Browder et al. (2009) maintained that students with SID are more likely to achieve these outcomes when they have regular opportunities to access literature (e.g., through adapted books), explicit instruction in how to access literature (e.g., instruction in listening comprehension), focused reading instruction (e.g., comprehension instruction), and opportunities to apply and generalize their reading skills (e.g., text applications). Browder and her colleagues (2009) also provided clear outcomes and instructional methods that educators could use to teach text comprehension to students with SID and emphasized that elementary literacy instruction for students with SID should focus on learning how to read (e.g., text comprehension), with SETs using evidence-based practices (EBPs), such as shared reading.

Evidence-Based Practice in Text Comprehension

Researchers have identified shared reading and systematic instruction as EBPs when providing literacy instruction to this population (Browder & Spooner, 2011), which are also effective when teaching text comprehension. Specifically, the identified include shared reading (Hudson & Test, 2011), and question answering (Browder, Wakeman, Spooner, Ahlgrim-Delzell, & Algozzine, 2006) for use during text comprehension instruction for students with SID. Additionally, prompting strategies, such as time delay (e.g., Browder, Trela, & Jimenez, 2007), simultaneous prompting (e.g., Alberto, Waugh, & Fredrick, 2010), and the system of least intrusive prompts (e.g., Browder, Lee, & Mims, 2011) have been used successfully, during text comprehension instruction for students with SID. However, in order to meet the unique learning needs of students with SID, the aforementioned practices should be chosen based on individual student needs, and SETs may need to use multiple, individualized strategies during instruction (i.e., multicomponent instruction).

Multicomponent Text Comprehension Instruction

Afacan, Wilkerson, and Ruppar (2017) defined multicomponent instruction as instruction that includes at least two of the five components of reading defined by the National Reading Panel (NRP) - phonemic awareness, phonics, fluency, vocabulary, and comprehension (2000). In their review, Afacan and colleagues (2017) found that when students with intellectual disability were exposed to multicomponent reading programs, they significantly improved their reading skills when compared with peers who received traditional sight word instruction. However, only seven studies were located, confirming a need for additional research investigating the impact of multicomponent instruction.

In this study, we defined multicomponent instruction as instruction that includes more than one of the individual component skills supporting text comprehension for students with SID (e.g., shared reading, question answering). One multicomponent text comprehension strategy that may meet the unique learning needs of students with SID is the...
CROWD (Whitehurst et al., 1994) in the CAR (Cole, Maddox, Lim, Yook, & Notari-Syverson, 2002) comprehension strategy (see Figure 1). CROWD (Completion, Recall, Open-ended, Wh questions, and Distancing prompts) helps adult readers remember what types of questions to ask during shared reading, which encourages adult-child interactions (Whitehurst et al., 1994). CAR (Comment and Wait, Ask Questions and Wait, and Respond by adding a little more) encourages adult-child interactions through wait time (Cole et al., 2002). Although no research has been conducted on these strategies with students with SID, when combined, the CROWD in the CAR comprehension strategy may offer SETs a teacher-led comprehension strategy based on multiple intervention components (i.e., question answering, shared reading) to teach text comprehension to students with SID, who perform at the early literacy level, and require intensive, explicit, contextualized instruction to improve text comprehension (see Browder & Spooner, 2011).

Professional Development Focused on Multicomponent Approaches

To keep up-to-date on effective literacy instruction, including multicomponent approaches to text comprehension for students with SID, SETs typically rely on professional development (PD). Unfortunately, PD often occurs outside of the classroom and infrequently aligns with ongoing teacher practice (Loucks-Horseley & Matsumoto, 1999). Although many SETs value literacy instruction for their students, they struggle when translating learned literacy content (e.g., text comprehension) into classroom practice. This often results in ineffective instruction for these students (Copeland & Keefe, 2004). However, when SETs receive PD that focuses on the learner, supports individual beliefs and values, and offers new skills or strategies for instruction (e.g., CROWD in the CAR comprehension strategy), teacher practice increases (McLeskey, 2011). Effective PD also includes active teacher learning, takes place over 20 or more hours, and includes job-embedded coaching (Croft, Coggshall, Dolan, Powers, & Killion, 2010; Desimone, 2009; Joyce & Showers, 1982).

Technology-enabled PD allows SETs to access effective PD that aligns with their learning needs and affords them multiple opportunities to practice using research-based instructional supports (Suppo & Mayton, 2014). For example, SETs can acquire new content knowledge, on demand, through virtual platforms, such as online modules, and receive job-embedded coaching online in real time (i.e., during instruction) aimed at transferring that knowledge to practice through advanced Bug-In-Ear (BIE) technology (Rock et al., 2009, 2012, 2014). Jimenez, Mims, and Baker (2016) investigated the use of online PD on SETs ability to acquire knowledge about mak-

Figure 1. CROWD in the CAR comprehension strategy.
ing data-based decisions. They found that while the online modules significantly increased teachers’ content knowledge, SETs teaching practices remained unchanged; thus, further underscoring the need for technology-enabled, real-time, job-embedded coaching, such as BIE coaching.

Once considered futuristic and costly, researchers across the US have demonstrated that BIE coaching, referred to hereafter as eCoaching, can be used effectively and efficiently to help pre- and in-service teachers and behavior analysts transfer new knowledge and skills into classroom practice (see Artman-Meeker, Rosenberg, Badgett, Yang, & Penney, 2017; Google, Rahn, & Ottley, 2015; Ottley & Hanline, 2014; Ploessl & Rock, 2014; Rock et al., 2009, 2012, 2014; Scheeler, McKinnon, & Stout, 2012). Moreover, social validity, among pre- and in-service SETs, for technology-enabled, real-time coaching is considered high (Ottley, Coogle, & Rahn, 2015; Ploessl & Rock, 2014; Rock et al., 2009, 2012). To date, however, what remains unknown are the effects of combining an online module with eCoaching when providing technology-enabled PD to SETs who teach learners with intensive support needs, such as those with SID.

The purpose of this study was to investigate the effects of PD delivered through an online module and eCoaching on SET’s use of a teacher-led comprehension strategy (i.e., CAR and CROWD comprehension strategy), during comprehension instruction for students with SID. This study builds on the aforementioned evidence base for teaching literacy to students with SID (e.g., Browder et al., 2011). Specifically, we extend the existing knowledge base by examining the use of a multicomponent text comprehension strategy (i.e., CROWD in the CAR comprehension strategy), providing technology-enabled PD (i.e., online module + eCoaching) to facilitate SETs use of this strategy, and examining comprehension and engagement outcomes for students with SID. To do so, we posed four research questions:

1) How does the online module plus eCoaching affect teachers’ implementation of the CROWD in the CAR comprehension strategy, as evidenced by opportunities to respond (OTR), during shared reading for students with SID?
2) In what ways does the online module plus eCoaching impact the frequency and types of questions asked during comprehension instruction when teachers use the CROWD in the CAR comprehension strategy, during shared reading for students with SID?
3) How does teachers’ use of the CROWD in the CAR comprehension strategy impact listening comprehension outcomes for students with SID (i.e., frequency and accuracy of responses)?
4) How does teachers’ use of the CROWD in the CAR comprehension strategy impact students with SID engagement in comprehension instruction?

Method

Participants

We used purposeful, convenience sampling to select teacher and student participants (Gall, Gall, & Borg, 2007). To be selected, SETs had to be providing literacy instruction to students with SID in the classroom, and students had to have a diagnosis of SID, which was defined as significant limitations in intellectual functioning and adaptive behavior (Spooner & Browder, 2011).

Teacher participants. Three Caucasian female SETs participated. SET participants ranged in age from 25–33 years and had 3–8 years of teaching experience. Additionally, levels of certification varied among the SETs. Emily had a Bachelor of Arts Degree in Psychology and was credentialed as a Registered Behavioral Technician (RBT). Carol was RBT certified, a Board-Certified Behavior Analyst, and had a degree in Business Administration. Courtney held an expired teaching license in special education, which she earned in another state. All three SETs were employed in a private school for children with SID and autism. None of the SETs had experience with the CROWD in the CAR comprehension strategy or eCoaching.

Student participants. Three Caucasian students with SID and autism participated in this study. Students ranged in age from 7–9 years. Anthony was a 9-year-old male diagnosed with SID, autism, cerebral palsy (CP), and hearing loss. He communicated through the use of
short (2-3 word) phrases and a picture exchange communication system (PECS). Anthony was classified as an emergent reader who attended to stories and pictures with prompts but was not yet recognizing words other than his own name. Behaviors that impeded Anthony’s daily instruction included screaming and aggression (e.g., hitting, kicking, biting). Also, Anthony was not yet toilet trained and began this process during the study. Mary was a 7-year-old female, with SID, autism, and speech delays. She communicated through scripting and short phrases. Mary was classified as an emergent reader who attended to texts for short periods of time with prompting, could recognize letters, was learning letter sounds, and was able to recognize her name in print. Mary often demonstrated stereotypic or repetitive behaviors and struggled with changes in her routine, which impeded her daily instruction. Jonathan was a 9-year-old male diagnosed with SID, autism, and attention deficit hyperactivity disorder (ADHD). He had difficulty verbalizing his wants and needs, communicated through scripting, and using short phrases when prompted by the teacher. Jonathan was classified as an emergent reader who was learning high frequency sight words. During instruction, Jonathan struggled with inattentiveness and often became upset by other students’ behaviors in the classroom. Due to confidentiality concerns expressed by administrators at the study site, additional identifying information regarding the students (e.g., IQ score, adaptive behavior scales) was not available. All students participated in daily literacy instruction.

Setting

The study took place in a private, separate school in the Southeastern United States that provides individualized diagnostic, therapeutic, and educational services to 335 children diagnosed with autism, intellectual disability, and other developmental disabilities. The school was staffed by three executive staff, six administrators, nine clinical staff, 39 behavioral technicians/teachers, and assistant teachers, volunteers, and student interns. Although we intended for literacy lessons to occur in each SETs classroom, during their regularly scheduled literacy lesson, technology barriers (i.e., inoperable, outdated computers) precluded some online connections. Consequently, each SET conducted their literacy lessons in Courtney’s classroom. The SETs carried out literacy lessons using one-on-one or small groups of two to five students. The first author provided eCoaching from a private, remote location (e.g., university/home office).

Materials and Equipment

During comprehension instruction, SETs used a storybook (e.g., The Gruffalo by Julia Donaldson) and an expository text (e.g., Apples by Gail Gibbons) while engaging in shared reading with students with SID. During the generalization phase, SETs read the expository text. To provide eCoaching during the intervention phase, we used a modified web-based interactive video conferencing system and advanced online bug-in-ear (BIE) system described in Rock et al. (2009). Specifically, the eCoach used the Internet, a web camera, and a MacBook Pro Laptop. SETs used the Internet, a computer, and a Plantronics VoIP USB Headset. All sessions were carried out via Skype and captured electronically, using Call Recorder for Mac v.2.5.16.

Independent Variable

The independent variable for this study was technology-enabled PD, which consisted of an online module plus eCoaching for the CROWD in the CAR comprehension strategy, where eCoaching was defined as “a relationship in which one or more persons’ effective teaching skills are intentionally and potentially enhanced through online or electronic interactions with another person” (Rock et al., 2014, p. 162).

Online module. SETs completed a self-paced online module which they accessed through a private, password protected Wiki-space account. The content of the online module reflected two out of the four components of effective PD (i.e., study of theory and peer observation of best practice) recommended by Joyce and Showers (1982). Through the online module, SETs gained content knowledge about the importance of effective literacy instruction for students with
SID, shared reading, and the CROWD in the CAR comprehension strategy. Additionally, SETs observed shared reading and the CROWD in the CAR comprehension strategy via online videos and were assessed at the completion of the online module through an eight-question assessment, using Survey Monkey.

ECoaching. SETs received one-on-one eCoaching from the eCoach (i.e., the first author) after successful completion of the online assessment. ECoaching reflected the third of four components of effective PD (i.e., one-on-one coaching) recommended by Joyce and Showers (1982). The fourth component (i.e., group coaching) was beyond the scope of this study. During the one-on-one eCoaching sessions, SETs were provided immediate feedback on their use of the CROWD in the CAR comprehension strategy during text comprehension instruction. The eCoach contacted SETs, at a pre-determined time, via Skype and provided feedback while the SET conducted the shared reading activity. Specifically, the eCoach encouraged SETs use of the CROWD in the CAR comprehension strategy (i.e., Excellent job asking a recall question!); instructed or corrected SETs use of the strategy (i.e., That was a yes or no question, try asking Where is the Gruffalo?), or asked clarifying questions (i.e., Did the Jonathan answer the previous question correctly?), in vivo. Since SETs were wearing a Plantronics VoIP USB Headset, they were able to hear the eCoach’s voice, but the students were not. All eCoaching sessions were scheduled in advance for no more than 15 minutes, which was the duration of shared reading.

Dependent Variables

There were two SET dependent variables: (1) SETs use of the CROWD in the CAR comprehension strategy, during shared reading for students with SID; and, (2) frequency and type of questions the SET asked when using the CROWD. The frequency of questions the teacher asked referred to the total amount of questions asked during the 15-minute shared reading session, specifically teacher-directed OTRs. The type of questions referred to CROWD (i.e., completion, recall, open-ended, WH, and distancing) and non-CROWD questions (e.g., Touch Gruffalo). Additionally, questions were categorized and coded as literal, inferential, critical, and evaluative comprehension questions, as well as higher order or lower order questions based on Blooms Taxonomy (see Anderson et al., 2001).

There were two student dependent variables used to assess the impact of the intervention on students’ listening comprehension and engagement: (1) number of correct, independent, responses to listening comprehension questions asked by the SET when using the CROWD in the CAR comprehension strategy; and (2) interval recordings of student engagement (i.e., total number of engaged intervals over the total possible intervals and reported as a percentage).

Experimental Design

We used a single subject, multiple-baseline across participants design (Gast, 2010) to investigate the effects of an online module plus eCoaching on SETs use of the CROWD in the CAR comprehension strategy during shared reading for students with SID and the impact of SETs’ use of this strategy on students’ listening comprehension. There were three SET/student dyads in this study, and each SET/student dyad entered baseline at the same time and were probed every day. Once a stable trend with at least five data points was obtained, the first dyad entered the intervention phase. The remaining participants continued in the baseline until they entered the intervention (Gast, 2010). Once the first student participant demonstrated a change in level and/or accelerating trend over the five data points, the second dyad entered the intervention. The same procedures were repeated for the last SET/student dyad.

Data Collection and Measures

As noted previously, all literacy lessons, during baseline and intervention phases, were recorded using Call Recorder for Mac. Due to the need to code multiple dependent variables simultaneously, videos were uploaded and stored on a secure and private database (i.e., password protected, encrypted, hard drive) for later data analysis, as well as IOA analysis.

SET measures. First, to measure SET’s use of the CROWD in the CAR comprehension
strategy, data were collected on mean rate of teacher provided opportunities to respond (OTR). OTRs provided by SETs included questions asked with the CROWD, but excluded requests or commands (i.e., “Touch the Gruffalo.”). Additionally, multiple related questions asked consecutively (i.e., cluster questions) were coded as one OTR. For example, “Where’s the Gruffalo? Who is this?” was counted as one OTR. Second, to measure the frequency and type of questions SETs asked when using the CROWD (i.e., Completion, Recall, Open-ended, WH-, and Distancing), data were collected on the number and type (e.g., literal, inferential, recall, or completion) of questions asked during the lesson.

Student measures. To measure the impact of SET’s use of the CROWD in the CAR comprehension strategy on students’ comprehension, data were collected on the frequency, accuracy, and type of questions answered by students. Although anecdotal notes were recorded on the type of question answered (e.g., literal, inferential, recall, or completion) and the type of response mode the student used (i.e., verbal or by touching the book), only independent, correct responses were graphed. Second, to measure student engagement data were collected via interval recording during 2-minute intervals during the literacy lesson (Cooper, Heron, & Heward, 2007). Specifically, student engagement was recorded as occurring (+) if the student was engaged at any time during each 2-minute interval across the duration of the lesson (i.e., 15 minutes; Cooper et al., 2007). Engagement was defined as the “student attending to (i.e., looking at) the teacher, making appropriate motor responses (e.g., following directions, manipulating materials), asking for assistance in an appropriate manner, and interacting with peers or adults within the structure of the activity” (Courtade, Lingo, & Whitney, 2013, p. 9).

Social Validity

At the end of the study, SETs completed a social validity questionnaire to assess the overall importance and feasibility of the study. SETs completed the survey via Qualtrics, an online platform used to design, develop, and analyze online surveys, comprised of ten 5-point Likert-type scale questions indicating their level of satisfaction with the intervention (i.e., online module plus eCoaching), as well as the feasibility and appropriateness of the CROWD in the CAR comprehension strategy with their students.

Treatment Fidelity

eCoaching. Rock and colleagues (2012) recommend eCoaches should provide four times (4x) as many encouraging (i.e., praise contingent after demonstrating a specific behavior) forms of feedback in relation to instructing (objective information), questioning (information seeking), or correcting (noting errors and providing specific ways to correct them). Therefore, we collected frequency data on the type of coaching statements provided during eCoaching via video recorded lessons. Then the first author and a trained observer examined and coded the frequency of the eCoach’s comments as encouraging, instructing/correcting, or questioning. Data indicated the eCoach did not meet the recommended 4:1 ratio discussed by Rock et al. (2012). Percentage of agreement for the 4:1 feedback ratio was calculated at 100%.

Teacher. To assess SET’s ability to implement shared reading and to give students opportunities to interact with the text a fidelity checklist for SET questioning with the CROWD was used. This checklist was modified from teacher task analyses used in previous research, but specific items related to the CROWD in the CAR comprehension strategy were added. Percentage of SET agreement to the 4-step checklist ranged from 33% to 100% across all phases.

Inter-Observer Agreement (IOA)

The first author and a second coder conducted IOA on 20% of all sessions across all phases (baseline, intervention, maintenance, and generalization; Gast, 2010) for each dependent variable (i.e., OTR, question frequency and level, student comprehension, and student engagement). The second coder had previous training and coursework in single subject research methodology, including coding as a form of data collection. To calculate reliability, the total number of agreements was divided by the total number of agreements plus...
disagreements, and multiplied that number by 100 (Cooper et al., 2007). IOA was calculated as 100% across all phases.

Procedure

Baseline. During baseline, SETs were instructed to read the storybook text (i.e., The Gruffalo) to their students as they typically would during the shared reading portion of a literacy lesson. During the lesson, data were collected on SET and student dependent variables (i.e., the first author recorded the data on data sheets).

Intervention. During the intervention phase, SETs completed the online module and received eCoaching on the CROWD in the CAR comprehension strategy during their literacy instruction. Immediately after meeting baseline criteria (i.e., 5 consecutive data points, stable or decelerating trend; Gast, 2010), the researcher sent the first participant, Emily, the web link via email to the online PD module. Once Emily completed the module, the researchers checked for completion and accuracy via the online quiz. Similar to baseline, Emily was instructed to carry out shared reading, using the storybook, during her scheduled literacy lesson, and to use the CROWD in the CAR comprehension strategy learned through the online PD module. During shared reading, Emily read the story and asked questions while reading the book to engage students with the text and to assess comprehension. Due to the nature of the CROWD in the CAR comprehension strategy, Emily was able to choose the types of questions she wanted to ask during the lesson (e.g., literal, inferential, recall, or completion). Additionally, the first author also provided eCoaching feedback (i.e., instructing/correcting, encouraging, or questioning) in situ, while Emily practiced using the CROWD in the CAR comprehension strategy during shared reading. The same procedure was repeated with the second and third SETs, Carol and Courtney, and correct responses to students were followed by verbal praise such as “Great job! The owl is in the tree.”

Carol and Courtney read the Gruffalo to Mary and Jonathan throughout the intervention phase. Emily introduced a new book (i.e., Pete the Cat by Eric Litwin) during session 14 because Anthony began to display problematic behaviors (i.e., screaming, yelling, refusal to answer questions) when sessions began; however, the chosen book was not comparable in length and text to the Gruffalo. Therefore, the first author instructed Emily to read If You Give a Dog a Donut by Laura Numeroff for the remainder of the intervention.

Maintenance and generalization. Maintenance and generalization data were recorded based on SETs use of the CROWD in the CAR comprehension strategy based on the aforementioned SET and student dependent variables. Once all SET/student dyad received the intervention and demonstrated growth as measured by a change in level and/or trend, they entered the maintenance phase simultaneously. During the maintenance phase, SETs were instructed to engage in shared reading with their students similar to the way they did during the intervention phase using the aforementioned storybooks. SETs and students were observed for no more than 15 minutes and no additional online module or eCoaching were provided.

Once all SETs demonstrated that they were able to maintain use of the CROWD in the CAR comprehension strategy without the module or eCoaching, they then immediately entered the generalization phase. During the generalization phase, SETs were instructed to engage in shared reading with their students as usual during scheduled literacy lesson, now using an expository text provided. The researcher observed the SETs and students for no more than 15 minutes, and consistent with the maintenance phase procedures, no additional online module or eCoaching were provided. Emily and Anthony did not complete the generalization phase because Emily took a position at another school prior to the conclusion of the study.

Results

Opportunities to respond (OTR). See Figure 2 for a visual representation of mean rate of OTR across all phases. Overall, each SET increased their mean rate of OTR in the intervention and maintenance phases; however, Carol and Courtney experienced an immediate increase from baseline to intervention. Emily increased her mean rate of OTR from
Figure 2. Mean rate for special education teacher opportunities to respond during instruction.
1.97 average in baseline to 2.99 average during intervention; Carol increased her mean rate of OTR from 0.39 average in baseline to 2.85 average during intervention; and Courtney increased her mean rate of OTR from 0.18 average in baseline to 2.54 average during intervention. All SETs also increased their mean rate of OTR from intervention to maintenance. Emily increased her mean rate of OTR from 2.85 during intervention to 3.04 in the maintenance phase. Finally, Courtney increased her mean rate of OTR from 2.54 during intervention to 5.00 in the maintenance phase. During generalization there was a decrease in the mean rate of OTR for all SETs. Emily decreased her mean rate of OTR from 3.15 in the maintenance phase to 2.90 during generalization, Carol decreased her mean rate of OTR from 3.04 in the maintenance phase to 2.40 during generalization, and Courtney decreased her mean rate of OTR from 5.00 in the maintenance phase 2.93 during generalization.

**Frequency and type of questions.** Emily asked an average of 27 questions during baseline (range = 20–33), 44 during intervention (range = 24–77), and 27 during maintenance (range = 26–28). Emily left the study during generalization and only one data point was collected. She asked 50 questions the generalization phase. Carol asked an average of 4 questions during baseline (range = 0–19), 24 during intervention (range = 13–17), 22 during maintenance (range = 9–17), and 16 during generalization (range = 8–12). Courtney asked an average of 1 question during baseline (range = 0–6), 28 during intervention (range = 19–37), 38 during maintenance (range = 28–45), and 36 during generalization (range = 20–48). See Table 1 for the frequency and types of questions across all phases.

**Impact on students’ listening comprehension.** Anthony answered an average 5% of his questions correctly during baseline, 24% during intervention, 47% in maintenance, and 28% during generalization. Mary answered an average 21%, 59%, 77%, and 70% correctly during baseline, intervention, maintenance, and generalization. During each respective phase, Jonathan answered an average 22%, 72%, 78%, and 34% correctly. See Figure 3 for a visual representation of listening comprehension outcomes for students across all phases.

**Student engagement.** Due to the nature of shared reading and because SETs read two highly engaging throughout the study, student engagement remained high. Student engagement for Anthony was 100% during baseline, maintenance, and generalization. During intervention, his engagement was 90%, ranging from 0% to 100. Engagement for Mary and Jonathan was 100% across all phases.

**Social validity.** Overall, all SETs agreed or strongly agreed the online module was accessible, practical, and useful. Additionally, all SETs agreed or strongly agreed the online module strengthened their skills as a teacher. They found the knowledge beneficial, and the CROWD in the CAR comprehension strategy easy to incorporate into their shared reading

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Emily</th>
<th>Carol</th>
<th>Courtney</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completion</td>
<td>44</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>Recall</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Open-Ended</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WH</td>
<td>64</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Distancing</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completion</td>
<td>249</td>
<td>157</td>
<td>84</td>
</tr>
<tr>
<td>Recall</td>
<td>8</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Open-Ended</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>WH</td>
<td>331</td>
<td>85</td>
<td>45</td>
</tr>
<tr>
<td>Distancing</td>
<td>35</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Maintenance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completion</td>
<td>43</td>
<td>61</td>
<td>139</td>
</tr>
<tr>
<td>Recall</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Open-Ended</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WH</td>
<td>84</td>
<td>51</td>
<td>53</td>
</tr>
<tr>
<td>Distancing</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Generalization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completion</td>
<td>8</td>
<td>2</td>
<td>86</td>
</tr>
<tr>
<td>Recall</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Open-Ended</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>WH</td>
<td>30</td>
<td>66</td>
<td>85</td>
</tr>
<tr>
<td>Distancing</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. WH questions include who, what, where, when, why, and how questions.
Figure 3. Number of independent correct responses to listening comprehension questions.
lessons. All SETs reported the eCoaching was not distracting during instruction and it increased their ability to use the CAR and CROWD during shared reading. Finally, two SETs reported that they noticed an increase in their students’ listening comprehension.

**Treatment fidelity.** The mean eCoaching fidelity for encouraging to instructing/correcting comments was 3.87:2.9 for Emily, 4.8:2.2 for Carol, and 4.8:4.6 for Courtney. Teacher fidelity was assessed by calculating the presence or absence of SETs reading in a shared reading format and if they asked questions using the CROWD in the CAR comprehension strategy. All SETs demonstrated 100% fidelity.

**Discussion**

Results of this study indicated that when provided an online module plus eCoaching, SETs learned a teacher-directed, multi-component comprehension strategy (i.e., the CROWD in the CAR), and carried it out effectively during shared reading for students with SID. As a result, students with SID experienced improved outcomes in comprehension, and a functional relationship was established between the teachers’ use of the CROWD in the CAR comprehension strategy and students’ independent correct comprehension responses during shared reading. These factors suggest that multicomponent comprehension instruction, online PD, and eCoaching may be effective for use with this population.

**Multicomponent Comprehension Instruction**

**Impact on listening comprehension for students with SID.** The variability within baseline data, especially for Mary and Jonathan, can be attributed to the dependent variable measurement (i.e., the amount of questions asked by SETs). The students’ OTRs varied, and this directly impacted the percentage of students’ correct responses as linked to the number of questions asked by SETs. For example, if the SET asked two questions and the student answered correctly both times, his or her percentage of independent correct responses for that session was 100%; however, if the SET asked 25 questions and the student responded correctly to 20 questions, then his or her percentage of independent correct responses for that session was 80%. Therefore, variable student data were linked to SET questioning. These results differ from those found by previous researchers in which comprehension questions were determined a priori (e.g., Browder et al., 2011) or were included in the teacher task analysis (e.g., Browder et al., 2007).

**Opportunities to respond (ORTs) during shared reading.** Overall, the students increased their accuracy of responses to the listening comprehension questions posed by the SETs, and although data were variable with some overlap between baseline and intervention, a functional relationship was found via an immediate change in level and trend in the students’ data suggesting that more OTRs provided during text comprehension instruction directly impacted their ability to answer the questions correctly. One contributing factor to an increase in students’ accuracy in responding may have been result of SETs using the CROWD in the CAR comprehension strategy and providing more OTRs. All SETs provided more OTRs during the intervention phase. Researchers recommend providing four to six responses per minute during the instruction of new materials with 80% accuracy, and eight to twelve responses per minute with 90% accuracy during independent practice (Sutherland, Alder, & Gunter, 2003); however, these guidelines refer to students without disabilities. Results of this study indicate that when teaching students with SID to comprehend text and to account for learning rate (see Copeland & Keefe, 2007), SETs may choose to reduce the amount of OTRs to about 3 per minute in order to accurately engage students in the text and to give them a chance to demonstrate their understanding, but more research is needed.

**Questioning.** Previous researchers who have investigated literacy and comprehension for this population taught students to answer literal comprehension questions (Hudson & Browder, 2014; Wood, Browder, & Flynn, 2015). In this study, the majority of questions asked were literal questions; however, when using the CROWD in the CAR comprehension strategy SETs asked more higher order (e.g., inferential) questions. Although no critical or evaluative questions were asked, results
indicate that using the CROWD in the CAR comprehension strategy may have prompted SETs to think more deeply about the types of questions they asked students with SID during text comprehension instruction. Over time, doing so may contribute to students with SID experiencing increased independence as readers (Browder et al., 2009).

**Online Module + eCoaching**

Previous researchers have investigated the use on online modules (Jimenez et al., 2016) and eCoaching (Artman et al., 2017; Ploessl & Rock, 2017; Rock et al. 2009, 2012, 2014) for pre- and in-service teachers and behavior analysts. However, this is the first study to investigate the use of an online module plus eCoaching with in-service SETs teaching students with intensive literacy needs, such as those with SID. Results of this study indicate that this combined approach was effective. Also, although not a component analysis, this study is the first to explore an online module plus eCoaching effectively to bolster SETs content knowledge of multicomponent comprehension instruction (i.e., CROWD in the CAR) and to facilitate that knowledge transfer into classroom practice; thus, expanding upon recommendations of Jimenez et al. (2016).

In order to be effective, Desimone (2009) recommended 20 hours of learning during PD; however, in this study, SETs received less than this amount of PD. Results indicated the intervention was effective in increasing SETs use the CROWD in the CAR comprehension strategy and in improving student’s comprehension outcomes, but the decelerating trend that occurred in generalization may be attributed to the length of time engaged in eCoaching. These findings are similar to those in which researchers removed eCoaching to test the participant’s ability to use the taught strategies independently (e.g., Ploessl & Rock, 2014) or during generalization (e.g., Google et al., 2015). Also, results of this study extend the findings previous research in which researchers provided online PD (e.g., Jimenez et al., 2016) or eCoaching (e.g., Rock et al., 2009), because in this study, SETs received both. By providing an alternative method in which SETs could learn a new literacy strategy (i.e., CROWD in the CAR) and providing eCoaching, researchers facilitated SETs content knowledge acquisition and transfer.

**Limitations and Implications for Future Research**

There were several limitations of this study. First, the setting of the research study was a limitation because the SETs shared the same technology throughout the study, and they were often in and out of each other’s classroom during their scheduled literacy times. We speculate that the close proximity and daily interactions may have slightly altered the way they read during their daily shared reading. Another limitation was that location of the study, which occurred in a separate school for students with SID. Future research should occur in a variety of settings (e.g., general education classroom) to test the feasibility of the intervention across contexts. Although texts are read repeatedly during literacy instruction and SETs posed a variety of questions, the length of time in which SETs read the text may be another limitation. Future researchers should investigate the use of the CROWD in the CAR comprehension strategy across multiple texts, expanding the generalization phase of this study. The independent variable was also a limitation, because researchers were unable to determine the impact of the CROWD in the CAR comprehension strategy from the impact of the module plus eCoaching. Future researchers should investigate the effects of these components individually to assess impact. Finally, future researchers should continue investigating the range of OTRs needed for students with SID.

**Conclusion**

In summary, the results of this study extended the prior research (e.g., Browder et al., 2011) on providing text comprehension instruction during shared reading for students with SID. Results indicated that SETs effectively implemented a multicomponent comprehension strategy (i.e., CROWD in the CAR) learned through technology-enabled professional development, consisting of an online module and eCoaching, which may provide a model that allows PD providers to provide support where it matters most – in and during the
transfer to classroom practice. Also, although by no means definitive, the results indicate that SETs use of a multicomponent comprehension instruction approach, namely the CROWD in the CAR, can lead to improved literacy outcomes for students with SID.

References


Erickson, K. A., Hansen, G., Hatch, P., & Sanders, E. (2009). Research-based practices for creating access to the general curriculum in reading and literacy for students with significant intellectual disabilities. Chapel Hill, NC: Center for Literacy and Disability Studies.


Hudson, M. E., & Test, D. W. (2011). Evaluating the evidence base of shared story reading to promote...
literacy for students with extensive support needs. *Research and Practice for Persons with Severe Disabilities, 36*, 34–45. doi:10.2511/rpseud.36.1-2.34


Received: 6 August 2018
Initial Acceptance: 3 October 2018
Final Acceptance: 23 November 2018
Systematic Review of Functional Communication Training in Early Care and Education Settings

Li Luo
Capital Normal University, Beijing

Nicholas A. Gage
University of Florida

Debra A. Prykanowski
Appalachian State University

Abstract: The purpose of this review was to quantitatively synthesize studies on functional communication training (FCT) for young children exhibiting challenging behavior when FCT was delivered in early care and education settings. Twenty studies published in peer-reviewed journals between 1992 and 2016 that met established inclusion criteria were reviewed and coded on variables associated with child-level, intervention-level, and study-level characteristics. The What Works Clearinghouse (WWC) standards for evaluating single-case design research were then applied to each study. Study effects were calculated using the percentage of non-overlapping data and Tau-U. Results indicated that FCT interventions had strong effects for reducing challenging behavior. Findings suggest additional studies with more children and methodological rigor are needed before FCT can be characterized as an evidence-based practice for the treatment of challenging behavior of young children with disabilities in early care and education settings based on the WWC’s evidence-based criteria. Limitations of the present synthesis and implications for future research are discussed.

The presence of severe and persistent challenging behavior has long-lasting consequences for young children, including peer rejection, academic failure, and social maladjustment (Dunlap et al., 2006; O’Connor, Dearing, & Collins, 2011). In the United States, the prevalence of challenging behavior in young children has been estimated to be between 10% and 30% (Fox, Dunlap, & Powell, 2002), and these percentages are even higher for young children with disabilities (Eisenhower, Baker, & Blacher, 2005). Many young children with or at risk for disabilities are at heightened risk for challenging behavior due to social interaction and communication skill delays. For example, research has shown that young children with developmental disabilities more often exhibit what are characterized as high-intensity challenging behaviors such as screaming, hitting, and biting than same-age typical peers (Sigafoos, 2000), and that up to 94.3% of children with autism spectrum disorders (ASD) display at least one type of challenging behavior as measured by the Autism Spectrum Disorder-Behavior Problems for Children (Matson, Wilkins, & Macken, 2009).

A critical need exists to identify evidence-based interventions to address challenging behaviors exhibited by children with or at risk for disabilities in early childhood settings (Conroy, Dunlap, Clarke, & Alter, 2005). A number of interventions have been developed and researched to address challenging behavior of young children with or at risk for disabilities in early childhood settings (Conroy, Dunlap, Clarke, & Alter, 2005). A number of interventions have been developed and researched to address challenging behavior of young children with or at risk for disabilities (Horner, Carr, Strain, Todd, & Reed, 2002), but few have received the attention garnered by functional communication training (FCT; Carr & Durand, 1985). Since its first published demonstration in the mid-1980s by Carr and Durand, the effects of FCT on reducing challenging behavior of young children...
has been studied over the ensuing three decades (Durand & Moskowitz, 2015).

**Functional Communication Training**

FCT is based on a single premise: challenging behavior often serves a communicative function (Durand & Moskowitz, 2015). FCT is a function-based, differential reinforcement procedure in which the individual is taught an appropriate communication response that serves the same function as the challenging behavior, but is a socially appropriate replacement behavior (Carr & Durand, 1985). Researchers have summarized a three-step process that FCT researchers and practitioners typically follow. First, a functional behavior assessment (FBA) is conducted that involves descriptive and/or experimental methods to identify the function of challenging behavior (e.g., attention, escape, tangible, sensory). Next, a communicative response is identified that matches the function of challenging behavior. This communicative response might be expressed using one of the following forms: verbal language, picture communication, gestures or signs, or assistive technology devices. Four criteria should be taken into account in selecting the communicative response: (a) a child’s capability of completing the response, (b) ease of teaching the response, (c) ease of understanding and acknowledging the response, and (d) the effectiveness/efficacy of the response in serving its identified function. Third, a treatment plan is designed to teach the communicative response. Implementation of the treatment plan usually involves prompting the communicative response, reinforcing the communicative response, placing the challenging behavior on extinction, shaping the communicative response, fading the prompts, or delaying the delivery of the reinforcement (see Figure 1; Mancil & Boman, 2010).

**Reviews of FCT Research**

To date, several narrative or quantitative FCT reviews have been conducted. Some reviews have focused on different features of the FCT literature, such as studies that explicitly programmed for generalization (Falcomata & Wacker, 2013) or reinforcement schedule thinning following FCT (Hagopian, Boelter, & Jarmolowicz, 2011). Some reviews have focused exclusively on children with a specific disability, such as ASD (Mancil, 2006), or have targeted practitioner-implemented FCT in classroom settings for school-aged children (e.g., Andzik, Cannella-Malone, & Sigafoos, 2016).
Kurtz, Boelter, Jarmolowicz, Chin, and Hagopian (2011) examined FCT as a treatment for challenging behavior displayed by individuals with intellectual disabilities. Twenty-eight single-case experimental design studies were included and, based on the results of those studies, the authors suggested there was sufficient empirical support to characterize FCT as a well-established treatment for a range of problem behaviors exhibited by individuals with intellectual disabilities. Although more than one-third of the participants in Kurtz and her colleagues’ (2011) review were classified as being in a birth through age 6 age category, only 16% of the participants received FCT in a school or educational setting (e.g., preschool), while the remaining individuals received FCT in either clinical, home, or community settings.

Durand and Moskowitz (2015) conducted a narrative review of 21 studies of FCT with young children 1 to 6 years of age. This review highlighted critical features of the FCT process and the important role of schedule thinning and contextual fit, while emphasizing some of the difficulties associated with conducting an FBA in early childhood education settings. Although the authors concluded that FCT is effective and can positively impact young children exhibiting challenging behavior, they did not (a) conduct a quantitative review of the research studies or (b) define FCT as an evidence-based practice for young children in early childhood education settings. The goal of the present study is to remedy this gap in the literature and evaluate the empirical support for FCT in early childhood education settings.

The primary purpose of the present systematic review was to examine the status and empirical rigor of the empirical evidence to date for FCT implemented in education settings (e.g., preschool) for young children with challenging behavior. We consider empirical rigor as a critical component following recommendations from evidence-based practice centers such as the U.S. Department of Education’s What Works Clearinghouse (2017), which sets forth guidelines for identifying high quality research to inform practice. A secondary purpose was to quantitatively synthesize the strength of the empirical evidence. Four research questions were used to guide this synthesis:

**Research Question 1:** What child characteristics and intervention features define the FCT studies included in this review?

**Research Question 2:** Using the What Works Clearinghouse’s (WWC) Pilot Single-Case Design Standards for evaluating single-case design evidence, does FCT conducted in early care and education settings have a sufficient evidence base to be defined as an evidence-based practice?

**Research Question 3:** To what extent does FCT conducted in early care and education settings decrease challenging behavior of young children with or at risk for disabilities?

**Research Question 4:** What is the pooled magnitude of the treatment effect on challenging behavior across the reviewed studies that met both WWC design and evidence standards?

**Method**

**Search Procedures**

PsychINFO, Education Full Text, ERIC (ProQuest), and Social Sciences Citation Index databases were searched using the terms *functional communication training* or *functional equivalence training* in October 2014, and electronic searches were updated in February 2017 to identify articles published since 2014. Limits were set for peer-reviewed journal articles written in English. No limits were placed on publication date. The abstract and method sections of each potential article were examined to determine if the study was eligible for inclusion. Five inclusion criteria were used. First, the study was an empirical study published in a peer-reviewed journal. Second, the study had to include at least one young child who was between the ages of 0 and 72 months at the beginning of the study. Third, the study had to be conducted in an early childhood education program or school. Studies in which FCT was conducted in outpatient/day treatment, clinical, inpatient, home, or community settings were deemed ineligible. Fourth, the study had to be conducted with young children demonstrating challenging behavior, and challenging behavior had to be targeted as a dependent variable. Fifth, the
independent variable in eligible studies had to be FCT (i.e., must name the intervention as FCT). Studies in which FCT was as a component of a comprehensive intervention package or program were excluded. Upon completion of the electronic search and screening, a forward and a backward search of all articles meeting inclusion criteria were conducted. Furthermore, reference lists of relevant FCT reviews were searched to locate additional studies not identified in the electronic searches. Literature search results were uploaded to EndNote X7 software.

Study Coding Procedures

To summarize the existing empirical literature on FCT studies in early care and education settings, an adapted version of the coding frame developed by Gage, Lewis, and Stichter (2012) was used to code child-level, intervention-level, and study-level characteristics of each study included in the review (see Table 1). As child participant and intervention features could vary within the same study, variables associated with child-level and intervention-level characteristics were coded individually for each young child participant. Consistent with the data extraction procedures conducted by Rakap, Snyder, and Pasia (2014), Biosoft’s UnGraph for Windows (Version 5.0) was used to recover the data from published graphs of included studies. Ungraph has been validated as a reliable and valid method of data extraction (e.g., intercoder agreement $r = .96$) and was recommended for conducting single-case meta-analysis (Shadish et al., 2009).

Application of What Works Clearinghouse Standards

All studies included were defined as single-case designs by the study authors. In order to assess the quality of the evidence base and to determine whether or not FCT could be characterized as an evidence-based practice for young children with challenging behavior in early care and education settings, we applied the WWC Pilot Single-Case Design Standards (What Works Clearinghouse, 2017) to all included studies.

Application of design standards. Each included study was reviewed using the study design rating criteria to determine whether the study (a) Meets Design Standards without Reservations, (b) Meets Design Standards with Reservations, or (c) Does Not Meet Design Standards. In order to meet WWC design standards, the following design criteria are required: (a) the independent variable (i.e., FCT) must have been systematically manipulated with the researcher determining when and how to change the independent variable conditions, (b) each outcome variable must be measured systematically over time and reported a measure of interobserver agreement (IOA) for no less than 20% of sessions that met minimum thresholds (e.g., averaged percentage agreement $\geq 0.8$, or Cohen’s Kappa $\geq 0.6$), (c) the study must include at least three attempts to demonstrate an intervention effect, and (d) phases must meet criteria involving the number of data points to quantify as an attempt to demonstrate an effect (What Works Clearinghouse, 2017).

Application of evidence standards. Following the application of the WWC design standards, each study that met design standards with or without reservations was reviewed using visual analysis procedures to determine if the study demonstrated a functional relationship between the independent variable (i.e., FCT) and the outcome variable (i.e., challenging behavior). WWC standards focus on six visual analysis features: level, trend, variability, immediacy of the effect, overlap, and consistency of data in similar phases, which were used to examine within- and between-phase data patterns. Four steps were used to rate each study: (a) assess baseline data pattern; (b) compare the level, trend, and variability for adjacent phases; (c) compare the overlap, immediacy of the effect, and consistency of data patterns in similar phases; and (d) combine the information from each of the phase comparisons to determine whether three demonstrations of an effect occur at different points in time. Based on these visual analyses procedures, studies were rated as providing either Strong Evidence, Moderate Evidence, or No evidence of a functional relationship (What Works Clearinghouse, 2017).

Application of replication standards. The WWC single-case design panel recommended
<table>
<thead>
<tr>
<th>Child-Level Characteristics</th>
<th>Intervention-Level Characteristics</th>
<th>Study-Level Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child age:</strong> Age was recorded in years or year/month as the author(s) presented in the studies.</td>
<td><strong>Functional behavioral assessment (FBA):</strong> Recorded whether the study authors conducted a functional assessment interview, functional assessment observation, functional analysis, or other method.</td>
<td><strong>Research design:</strong> Each single-case experimental design study was coded into one of the following categories: reversal/withdrawal, multiple baseline, alternating treatments, changing criterion, multiple probe, combined design, or other design.</td>
</tr>
<tr>
<td><strong>Child gender:</strong> Gender was recorded as a dichotomous male-female variable.</td>
<td><strong>Behavioral functions:</strong> Recorded the purpose of each child’s challenging behavior identified from FBA.</td>
<td><strong>Behavioral measure:</strong> Recorded the method to measure behavior change (e.g., direct observation, rating scales).</td>
</tr>
<tr>
<td><strong>Disability:</strong> The type(s) of disability of each child was recorded from descriptions in the reviewed studies.</td>
<td><strong>FBA agent:</strong> Recorded who conducted the FBA.</td>
<td><strong>Descriptive statistics:</strong> Recorded the basic statistics to present the observed behaviors that serve as FCT outcome (e.g., frequency, duration, probability, rate).</td>
</tr>
<tr>
<td><strong>Speech/language:</strong> Each child’s speech or language ability was recorded.</td>
<td><strong>FBA setting:</strong> Recorded the environment in which the FBA was implemented.</td>
<td><strong>Treatment fidelity:</strong> Recorded the presence or absence of treatment fidelity data collected, treatment fidelity measure used, who completed the measure, and fidelity data reporting.</td>
</tr>
<tr>
<td><strong>Challenging behavior:</strong> The operational definitions of challenging behaviors in each article was reviewed, which resulted in five categories of challenging behavior: aggression, noncompliance, elopement, self-injurious behaviors, and tantrums/disruptions.</td>
<td><strong>FCT agent:</strong> Recorded who conducted the FCT or taught the child replacement behavior (e.g., teacher, research, or therapist).</td>
<td><strong>Maintenance/generalization:</strong> Recorded the presence or absence of maintenance/generalization measured.</td>
</tr>
<tr>
<td><strong>School attended:</strong> Recorded the type of school each child was attending from descriptions in the reviewed study, such as special school, inclusive preschool, child care center, Head Start, and summer school program.</td>
<td><strong>FCT setting:</strong> Recorded the environment in which FCT was conducted or replacement behavior was taught (e.g., classroom, or tutoring room).</td>
<td><strong>Social validity:</strong> Recorded the presence or absence of social validity data collected, social validity measure used, who completed the measure, and results.</td>
</tr>
<tr>
<td><strong>Classroom attended:</strong> Recorded the type of classroom each child was attending from descriptions in the reviewed study, such as self-contained classroom, inclusive classroom, and general education classroom.</td>
<td><strong>Replacement behavior:</strong> Recorded the appropriate communication responses (that are functionally equivalent) to replace challenging behavior that each child was taught during FCT.</td>
<td></td>
</tr>
<tr>
<td><strong>Instruction in replacement behavior:</strong> Recorded whether children were taught replacement behavior prior to the FCT intervention phase(s), and the length/duration of the instruction in replacement behavior.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the “5-3-20” threshold for combining the findings of single-case design studies into a single summary rating (Kratochwill et al., 2013). In other words, in order to be classified as an evidence-based practice, the following thresholds must be met by a body of single-case design studies in which strong or moderate evidence of a functional relationship was found: (a) there is a minimum of five single-case design studies examining the intervention that were rated to meet both design and evidence standards, (b) the single-case design studies must be conducted by at least three different research teams with no overlapping authorship at three different institutions, and (c) the combined number of cases (i.e., participants, classrooms, etc.) totals at least 20 (What Works Clearinghouse, 2017).

Magnitude of Treatment Effect Estimates

There is no universal agreement on effect size metrics in the evaluation of single-case design studies. Nonparametric, nonoverlap methods have been widely used in the quantitative syntheses of single-case design studies (Maggin, O’Keefe, & Johnson, 2011). Shadish, Hedges, Horner, and Odom (2015) have advocated for the use of standardized effect sizes, which they define as “statistics that describe the magnitude of an effect on a common scale” (p. 27). For ease of comparison and interpretation, we chose nonoverlap effect size metrics for included studies. It has been suggested that recommended practice in single-case design meta-analysis is calculating more than one effect size metric (Brossart, Vannest, Davis, & Patience, 2014). Therefore, two widely used nonoverlap effect size metrics (Gage & Lewis, 2013) were used: percentage of nonoverlapping data (PND) and Tau-U. PND is determined by calculating the percentage of successive intervention data points that do not overlap with the baseline data points. Since ABAB or reversal/withdrawal designs have two A-B phases, the combined nonoverlapping data across the two AB phases were calculated in the effect size estimation. For example, if 7 of 9 data points from the first B phase, and 5 of 7 data points from the second B phase did not overlap data observed in the baseline phases, PND was calculated as the formula: (7 of 9) + (5 of 7) = [(7+5)/(9+7)]*100% = (12/16)*100% = 75% (Scruggs & Mastropieri, 2013). If a study involved more than one young child, PND for that study was aggregated by taking the median as the study estimate.

Tau-U is an index that combines nonoverlap between phases with trend from within the intervention phase (Parker, Vannest, Davis, & Sauber, 2011). With the ability to control for trend, Tau-U is the percentage of data that demonstrate improvement over time by comparing all data points in each phase (Parker et al., 2011). Taken into account all data points in each phase, Tau-U is less influenced by outliers and a small number of data points. Tau-U was calculated using a web-based calculator (http://www.singlecaseresearch.org). If an ABAB or reversal/withdrawal design study included more than one young child, Tau-U was calculated and presented separately for each child.

Meta-Analytic Procedures

The single-case design research has not been widely used in evaluating the empirical evidence base to document evidence-based practices mainly due to the lack of agreed upon statistical methods for the analysis and meta-analysis of single-case design studies (Shadish, 2014). Through the comparison of seven effect size metrics in single-case design studies, Chen and her colleagues (2016) found some effect size metrics are comparable or interchangeable, which further supports the use of meta-analyses of single-case design studies. Recently, modern meta-analytic methods involving standardized effect size measures have been recommended for use to conduct, interpret, and summarize single-case design research (Shadish et al., 2015). To date, two meta-analyses of single-case design studies (i.e., Chezan, Wolfe, & Drasgow, 2017; Heath, Ganz, Parker, Burke, & Ninci, 2015) have been identified to examine the overall effectiveness of FCT on challenging behavior of individuals with disabilities, in which the magnitude of effect was measured for individual studies using Tau-U and robust improvement rate difference respectively.

For the present study, Tau-U was used in the meta-analysis. To calculate an omnibus effect-size for studies meeting WWC standards, we
employed a random-effects meta-analysis model. A random-effects model was chosen because: (a) the included studies were similar, but not functionally equivalent (for example, the characteristics of child participants and interventions in these studies differed in ways that could impact the distribution of effect sizes); and (b) the goal of the effect size is to generalize to a larger population (Raudenbush, 2009).

The meta-analysis was conducted using the metafor package (Viechtbauer, 2010) in R (R Development Core Team, 2012) and included only studies receiving a rating of Strong Evidence or Moderate Evidence based on WWC visual analysis procedures. The metafor package is suggested as an appropriate software for conducting meta-analyses for single-case design research (Shadish et al., 2015).

Reliability of Coding

All coded variables and effect size estimates were double coded and all discrepancies were resolved through mediation and, if necessary, a third opinion. The two primary coders for data extraction and ES estimates were doctoral students majored in early childhood special education who had completed doctoral-level specialized coursework on single-case designs and systematic review. For the application of the WWC standards, three independent reviews were conducted. The third reviewer is a WWC-certified single-case design reviewer and led all WWC standards mediation sessions.

Results

A total of 634 records were located through the systematic search of electronic databases and ancestral search. A PRISMA flow diagram of study selection is shown in Figure 2. After title and abstract screening, 165 records were retained for full-text review. Eighteen journal articles describing 20 studies met the inclusion criteria and were included in this systematic review.

Child-Level Characteristics

Thirty-six young children were identified in the 20 included studies. Table 2 provides descriptive information of participants and intervention characteristics by child. Almost all studies included in the present review provided child demographic information, including age, gender, disability, speech/language ability, and challenging behavior. Twenty-nine of the participating children were boys (81%) and seven were girls (19%). The children ranged in age from 2 years and 9 months to 6 years, with an average age being 55 months. All children were identified as having a disability or multiple disabilities. Forty-two percent of the children were reported as having ASD (n = 15), followed by intellectual disability (33%) and developmental delay (22%). More than 70% of these young children were enrolled in a special education school/classroom for children with disabilities. The majority of children were identified as nonverbal or demonstrated limited expressive language, and communicated primarily with gestures, pointing, or using one- to three-word utterances.

Challenging behavior varied among participants, but most children exhibited more than one category of challenging behavior. Tantrum/disruptive behavior and aggression were among the most frequently occurred categories of challenging behavior reported by either parents or teachers. Although most of the studies operationally defined children’s challenging behavior, definitions of challenging behavior for seven children in two included studies were not provided.

Intervention-Level Characteristics

FBA was used to assess the function of challenging behavior in almost all included studies, but the procedures varied across studies. Fourteen studies (70%) involving 28 young children used functional analysis (i.e., experimental manipulations). The functions of children’s challenging behavior included: access to adult attention (36%), access to preferred items/activities (28%), escaping demands/tasks (19%), avoiding transitions to unpredictable activities (3%), and multiple functions that combined more than one function (14%). After the functions of challenging behavior were identified, children were taught a communicative response to replace their challenging behavior. Communicative responses
fit into one of the following forms: verbal language (33%), picture communication (28%), assistive technology devices (14%), gestures/sign language (8%), and combined responses that involved more than one form of communicative response (17%).

Researchers (47%) were the primary FBA agent across children, while other agents included classroom teachers (25%), teams (i.e., involving more than one type of professional; 22%), and non-specified (6%). More than one third of the children (39%) were assessed in their regular classrooms to identify the function of their challenging behavior, while other FBA settings included a treatment room, a tutorial room, a room adjacent to the classroom, and an unused room in schools. Classroom teachers (42%) served as the primary individuals who taught the children communicative responses, other FCT agents included researchers (25%), teams (19%), therapists (3%), and non-specified (11%). More than half of the children (56%) were taught communicative responses in their regular classrooms.

Fourteen young children in five studies were taught communicative responses prior to the FCT sessions in which intervention data were collected, that is, instruction in communicative responses occurred between baseline and intervention conditions. However, there was great variability with respect to the frequency and duration of instruction in communicative responses across studies. Table 3 presents a summary of instruction in communicative responses and study-level characteristics by each study.
### TABLE 2
Summary of Child Participants and Interventions Characteristics in FCT Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Age</th>
<th>G</th>
<th>Speech/Language</th>
<th>School (Classroom)</th>
<th>CR</th>
<th>FRA</th>
<th>Functions</th>
<th>FBA Agent</th>
<th>FBA Setting</th>
<th>FCT Agent</th>
<th>FCT Setting</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis et al. (2009)</td>
<td>1</td>
<td>4y</td>
<td>M</td>
<td>ASD &amp; DLD</td>
<td>1- to 3-word utterances</td>
<td>Private school for children with ASD</td>
<td>Agg</td>
<td>NR</td>
<td>Access to tangible</td>
<td>N/A</td>
<td>N/A</td>
<td>Therapist</td>
<td>Therapy room</td>
</tr>
<tr>
<td>Davis et al. (2014)</td>
<td>1</td>
<td>6y</td>
<td>M</td>
<td>ASD &amp; ID</td>
<td>1- to 3-word utterances</td>
<td>Public school (Self-contained classroom)</td>
<td>TD</td>
<td>FA</td>
<td>Access to tangible</td>
<td>Researcher</td>
<td>Classroom</td>
<td>Researcher</td>
<td>Classroom</td>
</tr>
<tr>
<td>Durand &amp; Carr (1992)</td>
<td>1-6</td>
<td>4y8m</td>
<td>F</td>
<td>ASD</td>
<td>22 months (GELS)</td>
<td>School for children with DD</td>
<td>TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Room</td>
<td>Researcher</td>
<td>Room</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5y2m</td>
<td>M</td>
<td>MR</td>
<td>NR</td>
<td>TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Room</td>
<td>Researcher</td>
<td>Room</td>
<td>Verbal language</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5y6m</td>
<td>M</td>
<td>DLD</td>
<td>NR</td>
<td>TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Room</td>
<td>Researcher</td>
<td>Room</td>
<td>Verbal language</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4y11m</td>
<td>M</td>
<td>MR</td>
<td>NR</td>
<td>TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Room</td>
<td>Researcher</td>
<td>Room</td>
<td>Verbal language</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5y6m</td>
<td>M</td>
<td>DLD</td>
<td>NR</td>
<td>TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Room</td>
<td>Researcher</td>
<td>Room</td>
<td>Verbal language</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>5y</td>
<td>M</td>
<td>DLD</td>
<td>NR</td>
<td>TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Room</td>
<td>Researcher</td>
<td>Room</td>
<td>Verbal language</td>
</tr>
<tr>
<td>Durand (1993)</td>
<td>1</td>
<td>5y6m</td>
<td>F</td>
<td>CP &amp; MR</td>
<td>Yes (NS)</td>
<td>18 months (GELS)</td>
<td>Agg</td>
<td>TD</td>
<td>MAS</td>
<td>Multiple</td>
<td>Teacher N/A</td>
<td>Teacher Classroom</td>
<td>VOCD</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5y6m</td>
<td>M</td>
<td>MR</td>
<td>Yes (NS)</td>
<td>15 months (GELS)</td>
<td>Agg</td>
<td>TD</td>
<td>MAS</td>
<td>Access to tangible</td>
<td>Teacher N/A</td>
<td>Teacher Classroom</td>
<td>VOCD</td>
</tr>
<tr>
<td>Durand (1999)</td>
<td>1</td>
<td>5y6m</td>
<td>M</td>
<td>CP &amp; MR</td>
<td>NR</td>
<td>18 months (GELS)</td>
<td>Agg</td>
<td>TD</td>
<td>MAS</td>
<td>Escape task</td>
<td>Teacher Classroom</td>
<td>Teacher Classroom</td>
<td>ACD</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5y6m</td>
<td>M</td>
<td>MR</td>
<td>NR</td>
<td>15 months (GELS)</td>
<td>Agg</td>
<td>TD</td>
<td>MAS</td>
<td>Teacher Classroom</td>
<td>Teacher Classroom</td>
<td>Teacher Classroom</td>
<td>ACD</td>
</tr>
<tr>
<td>Fragale et al. (2010)</td>
<td>1</td>
<td>4y</td>
<td>M</td>
<td>ASD &amp; CARS</td>
<td>Short phrases or sentences</td>
<td>NR (Self-contained classroom)</td>
<td>TD</td>
<td>FA</td>
<td>Access to tangible</td>
<td>Researcher</td>
<td>NR</td>
<td>Teacher Classroom</td>
<td>Verbal requests</td>
</tr>
<tr>
<td>Gibson et al. (2010)</td>
<td>1</td>
<td>4y</td>
<td>M</td>
<td>Autism</td>
<td>Limited</td>
<td>E Interv. &amp; FA</td>
<td>Access to tangible</td>
<td>Researcher Classroom</td>
<td>Teacher Classroom</td>
<td>Hand rising</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hines &amp; Simonsen (2008)</td>
<td>1</td>
<td>3y6m</td>
<td>M</td>
<td>Autism</td>
<td>Nonverbal</td>
<td>E Interv. &amp; FA</td>
<td>Access to tangible</td>
<td>Researcher Classroom</td>
<td>Paraprofessional &amp; Researcher Classroom</td>
<td>Picture cards</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Age</th>
<th>G</th>
<th>Dx.</th>
<th>Tool</th>
<th>Speech/Language</th>
<th>School (Classroom)</th>
<th>CB</th>
<th>FBA</th>
<th>Functions</th>
<th>FBA Agent</th>
<th>FBA Setting</th>
<th>FCT Agent</th>
<th>FCT Setting</th>
<th>RB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambert et al. (2012)</td>
<td>1</td>
<td>5–6 y</td>
<td>F</td>
<td>DD</td>
<td>NR</td>
<td>One-syllable words</td>
<td>Preschool (Self-contained classroom)</td>
<td>Agg</td>
<td>FA</td>
<td>Multiple</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Teacher</td>
<td>Classroom</td>
<td>PBCards</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5–6 y</td>
<td>F</td>
<td>DD</td>
<td>NR</td>
<td>One-word phrases</td>
<td>Preschool (Self-contained classroom)</td>
<td>TD</td>
<td>FA</td>
<td>Multiple</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Teacher</td>
<td>Classroom</td>
<td>PBCards</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5–6 y</td>
<td>M</td>
<td>DD</td>
<td>NR</td>
<td>Nonverbal</td>
<td>Preschool (Self-contained classroom)</td>
<td>Agg</td>
<td>FA</td>
<td>Access to attention</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Teacher</td>
<td>Classroom</td>
<td>PBCards</td>
</tr>
<tr>
<td>Langdon et al. (2008)</td>
<td>1</td>
<td>6 y</td>
<td>M</td>
<td>Autism &amp; MR</td>
<td>SIT</td>
<td>Nonverbal</td>
<td>School for children with DD</td>
<td>SB</td>
<td>Interv.</td>
<td>Escape demand</td>
<td>Teacher</td>
<td>N/A</td>
<td>Researcher</td>
<td>Tutoring room</td>
<td>PECS</td>
</tr>
<tr>
<td>Najdowski et al. (2008)</td>
<td>1</td>
<td>5 y</td>
<td>M</td>
<td>DD</td>
<td>NR</td>
<td>NR</td>
<td>Kindergarten</td>
<td>TD</td>
<td>FA</td>
<td>Access to tangible</td>
<td>Researcher</td>
<td>Treatment room</td>
<td>Researcher</td>
<td>Treatment room</td>
<td>Verbal language teach</td>
</tr>
<tr>
<td>O’Neill &amp; Sweetland-Baker (2001)</td>
<td>1</td>
<td>6 y</td>
<td>M</td>
<td>Autism &amp; MR</td>
<td>DSM-IV &amp; AAMR</td>
<td>No functional speech</td>
<td>Elementary school (Self-contained classroom)</td>
<td>TD</td>
<td>FA</td>
<td>Escape demand</td>
<td>NR</td>
<td>Classroom</td>
<td>NR</td>
<td>Classroom</td>
<td>Sign or touch card</td>
</tr>
<tr>
<td>Peck et al. (2000)</td>
<td>1</td>
<td>5 y</td>
<td>M</td>
<td>AS &amp; ID</td>
<td>CT</td>
<td>18 months (VZ)</td>
<td>Daycare facility for children with ID</td>
<td>Agg &amp; TD</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher &amp; Caretaker</td>
<td>Room</td>
<td>Researcher &amp; Caretaker</td>
<td>Room</td>
<td>PECS</td>
</tr>
<tr>
<td>Radstaake et al. (2012)</td>
<td>2</td>
<td>5 y</td>
<td>F</td>
<td>AS &amp; ID</td>
<td>CT</td>
<td>26 months (VZ)</td>
<td>Daycare facility for children with ID</td>
<td>TD</td>
<td>Agg</td>
<td>Access to attention</td>
<td>Researcher &amp; Caretaker</td>
<td>Room</td>
<td>Researcher &amp; Caretaker</td>
<td>Room</td>
<td>PECS</td>
</tr>
<tr>
<td>Radstaake et al. (2013)</td>
<td>3</td>
<td>5 y</td>
<td>M</td>
<td>AS &amp; ID</td>
<td>CT</td>
<td>26 months (VZ)</td>
<td>Daycare facility for children with ID</td>
<td>Agg</td>
<td>FA</td>
<td>Access to tangible</td>
<td>Caretaker</td>
<td>Classroom</td>
<td>Caretaker</td>
<td>Classroom</td>
<td>PECS</td>
</tr>
<tr>
<td>Radstaake et al. (2015)</td>
<td>1</td>
<td>6 y</td>
<td>M</td>
<td>AS &amp; ID</td>
<td>CT</td>
<td>Nonverbal</td>
<td>Specialized daycare center</td>
<td>SIB &amp; TD</td>
<td>AA &amp; FA</td>
<td>Escape task</td>
<td>Teacher &amp; Researcher</td>
<td>Classroom</td>
<td>Teacher &amp; Researcher</td>
<td>Classroom</td>
<td>SGD</td>
</tr>
<tr>
<td>Reeve &amp; Carr (2000)</td>
<td>1</td>
<td>5 y</td>
<td>M</td>
<td>Multiply</td>
<td>NR</td>
<td>NR</td>
<td>Early intervention center</td>
<td>NS</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Tutorial classroom</td>
<td>Researcher &amp; Teacher</td>
<td>Tutorial classroom</td>
<td>Verbal or nonverbal requests</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5 y</td>
<td>M</td>
<td>SI</td>
<td>NR</td>
<td>NR</td>
<td>Early intervention center</td>
<td>NS</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Tutorial classroom</td>
<td>Researcher &amp; Teacher</td>
<td>Tutorial classroom</td>
<td>Verbal or nonverbal requests</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.5 y</td>
<td>M</td>
<td>SI</td>
<td>NR</td>
<td>NR</td>
<td>Child development center</td>
<td>NS</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Classroom</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Verbal or nonverbal requests</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.5 y</td>
<td>F</td>
<td>PPD</td>
<td>NR</td>
<td>NR</td>
<td>Early intervention center</td>
<td>NS</td>
<td>FA</td>
<td>Access to attention</td>
<td>Researcher</td>
<td>Tutorial classroom</td>
<td>Researcher &amp; Teacher</td>
<td>Tutorial classroom</td>
<td>Verbal or nonverbal requests</td>
</tr>
</tbody>
</table>

(Continued)
TABLE 2—(Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Age</th>
<th>G</th>
<th>Dis.</th>
<th>Tool</th>
<th>School (Classroom)</th>
<th>CB</th>
<th>FBA</th>
<th>Functions</th>
<th>FBA Agent</th>
<th>FBA Setting</th>
<th>FCT Agent</th>
<th>FCT Setting</th>
<th>RB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schindler &amp; Horner (2005)</td>
<td>1</td>
<td>5y</td>
<td>M</td>
<td>Autism &amp; CS</td>
<td>One-word utterances</td>
<td>Early intervention program</td>
<td>Agg &amp; TD</td>
<td>Interv.</td>
<td>Access to tangible</td>
<td>Parent &amp; Teacher</td>
<td>N/A</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Pointing</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4y</td>
<td>F</td>
<td>Autism</td>
<td>Vocabulary of 2 – 5 words</td>
<td>Early intervention program</td>
<td>TD</td>
<td>Interv.</td>
<td>Escape task</td>
<td>Parent &amp; Teacher</td>
<td>N/A</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Pointing &amp; verbal language</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4y</td>
<td>M</td>
<td>Autism</td>
<td>1- to 2-word utterances</td>
<td>Preschool</td>
<td>TD &amp; N</td>
<td>Interv.</td>
<td>Avoid transition</td>
<td>Parent &amp; Teacher</td>
<td>N/A</td>
<td>Teacher</td>
<td>Classroom</td>
<td>Use of cards</td>
</tr>
<tr>
<td>Volkert et al. (2009)</td>
<td>1</td>
<td>5y</td>
<td>M</td>
<td>Autism or DD</td>
<td>Extensive</td>
<td>Public school (Self-contained classroom)</td>
<td>Agg &amp; TD</td>
<td>Interv.</td>
<td>Multiple</td>
<td>Researcher</td>
<td>Unused room</td>
<td>NR</td>
<td>Verbal language</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5y</td>
<td>M</td>
<td>Autism or DD</td>
<td>One-word utterances</td>
<td>Public school (Self-contained classroom)</td>
<td>Agg &amp; TD</td>
<td>Interv.</td>
<td>Escape demand</td>
<td>Researcher</td>
<td>Unused room</td>
<td>NR</td>
<td>Sign language</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5y</td>
<td>M</td>
<td>Autism or DD</td>
<td>4- to 5-word utterances</td>
<td>Prekindergarten program</td>
<td>Agg</td>
<td>Interv.</td>
<td>Multiple</td>
<td>Researcher</td>
<td>Unused room</td>
<td>NR</td>
<td>Verbal language</td>
<td></td>
</tr>
</tbody>
</table>

Note. n = number of young child participant(s); G = gender; Dis. = disability; CB = challenging behavior; FBA = functional behavior assessment; FCT = functional communication training; RB = replacement behavior; M = male; F = female; N/A = not applicable; NR = not reported.

Disability: ID = intellectual disability, MR = mental retardation; CP = cerebral palsy; DD = developmental delays; AS = Angelman syndrome; CS = Charge syndrome; DLD = developmental language disorder; TS = tuberous sclerosis; HI = hearing impaired; PDD = Pervasive Developmental Disorder-Not Otherwise Specified; SI = speech impaired; DD = developmental disabilities (not specified); GS = Goltz syndrome; VI = visual impairment.

Tool: CARS = Childhood Autism Rating Scale (Schopler & Reichler, 1980); VABS = Vineland Adaptive Behavior Scales (Sparrow et al., 1984); SIT = Slosson Intelligence Test; CT = Chromosomal testing.


School/Cl.: SEC = special education classroom, GEC = general education classroom, IEC = inclusive education classroom.

CB: Agg. = aggression, N = noncompliance, E = elopement, SIB = self-injurious behavior, TD = tantrums/disruption.

FBA: Interv. = interview, ABC = antecedent-behavior-consequence observation, MAS = Motivation Assessment Scale, FA = functional analysis, AA = antecedent analysis.

FBA setting: Room = room adjacent to the classroom.

RB: PECS = Picture Exchange Communication System, ACD = assistive communication devices, VOCD = voice output communication device, SGD = speech generating devices.
<table>
<thead>
<tr>
<th>Study</th>
<th>INS in RB (Duration)</th>
<th>Outcomes</th>
<th>Research Design</th>
<th>DM</th>
<th>TF</th>
<th>M/G</th>
<th>SV</th>
<th>WWC Rating</th>
<th>Visual Analysis</th>
<th>PND for CB</th>
<th>Tau-U for CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis (2009)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>A-B design</td>
<td>POI</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Davis (2014)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>A-B design</td>
<td>POI</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Durand (1992)</td>
<td>Yes (5-45 min)</td>
<td>CB &amp; RB*</td>
<td>Multiple baseline</td>
<td>POI</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Meet with R</td>
<td>Moderate Evidence 100.00%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Durand (1993)</td>
<td>Yes (2-3 weeks)</td>
<td>CB, RB*, &amp; PFE</td>
<td>Multiple baseline</td>
<td>POI</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Meet with R</td>
<td>No Evidence    93.94%</td>
<td>—</td>
<td>—0.93</td>
</tr>
<tr>
<td>Durand (1999)</td>
<td>Yes (10-19 days)</td>
<td>CB &amp; RB</td>
<td>Multiple baseline</td>
<td>POI</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Meet with R</td>
<td>No Evidence    92.86%</td>
<td>—0.99</td>
<td></td>
</tr>
<tr>
<td>Fragale (2016)</td>
<td>NR</td>
<td>CB &amp; AE</td>
<td>Alternating treatment</td>
<td>POI</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Gibson (2010)</td>
<td>NR</td>
<td>CB</td>
<td>Reversal</td>
<td>POI</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Meet with R</td>
<td>Strong Evidence 100.00%</td>
<td>—1.00</td>
<td></td>
</tr>
<tr>
<td>Hines (2008)_Study 1</td>
<td>NR</td>
<td>CB &amp; AEB</td>
<td>A-B design</td>
<td>POI</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Hines (2008)_Study 2</td>
<td>NR</td>
<td>RB</td>
<td>A-B-C design</td>
<td>Frequency</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lambert (2012)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>Multiple baseline</td>
<td>Rate</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Meet with R</td>
<td>No Evidence    93.33%</td>
<td>—0.95</td>
<td></td>
</tr>
<tr>
<td>Langdon (2008)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>B-A-B design</td>
<td>Frequency</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Najdowski (2008)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>Reversal</td>
<td>Rate</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>100.00%</td>
<td>—1.00</td>
</tr>
<tr>
<td>O’Neill (2001)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>Multiple probe</td>
<td>POI</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>67.50%</td>
<td>—0.73</td>
</tr>
<tr>
<td>Peck (1996)</td>
<td>Yes (10 to 30 min per day)</td>
<td>CB &amp; RB*</td>
<td>Combined</td>
<td>Frequency</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Does Not Meet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Radstaake (2012)</td>
<td>NR</td>
<td>CB</td>
<td>Reversal</td>
<td>Frequency</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Does Not Meet</td>
<td>25.46%</td>
<td>—0.52</td>
<td></td>
</tr>
<tr>
<td>Radstaake (2013)*</td>
<td>NR</td>
<td>CB &amp; RB*</td>
<td>Reversal</td>
<td>Frequency</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Meet with R</td>
<td>Strong Evidence 83.33%</td>
<td>—0.98</td>
<td></td>
</tr>
<tr>
<td>Reeve (2000)</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>A-B design</td>
<td>POI</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Does Not Meet</td>
<td>—</td>
<td>75.00%</td>
<td>—0.58</td>
</tr>
<tr>
<td>Schindler (2005)</td>
<td>Yes (30-50 min)</td>
<td>CB &amp; RB</td>
<td>Multiple baseline</td>
<td>POI</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Does Not Meet</td>
<td>—</td>
<td>50.00%</td>
<td>—0.68</td>
</tr>
<tr>
<td>Volker (2009)_Study 1</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>Reversal</td>
<td>Rate</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Meet with R</td>
<td>No Evidence    33.33%</td>
<td>—0.81</td>
<td></td>
</tr>
<tr>
<td>Volker (2009)_Study 2*</td>
<td>NR</td>
<td>CB &amp; RB</td>
<td>Reversal</td>
<td>Rate</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Meet with R</td>
<td>No Evidence    44.83%</td>
<td>—0.23</td>
<td></td>
</tr>
</tbody>
</table>

Note. ES = effect size; INS in RB = instruction in replacement behavior; DM = descriptive statistics of dependent variable measure; TF = treatment fidelity; M/G = maintenance or generalization; SV = social validity; CB = challenging behavior; RB = replacement behavior; NR = Not reported; PFE = positive facial expression; AEB = appropriately engaged behavior; AE = academic engagement; Y = be assessed; N = not be assessed; POI = percentage of intervals; Meet with R = Meets WWC Pilot Single-Case Design Standards with Reservations; Does Not Meet = Does Not Meet WWC Pilot Single-Case Design Standards.

a Subjects’ graphs evaluated separately because they are ABAB designs.
b Only two young children in sample, therefore, functional relationship could not be established for the target population. * no data on replacement behavior were collected in baseline(s).
Study-Level Characteristics

Five studies included were, after careful review, determined to be A-B designs or A-B-C designs, while the remaining 15 studies employed single-case experimental designs. The average number of young children per study was 2 (range = 1 to 6). Across these 20 included studies, the most frequently used single-case experimental design was reversal/withdrawal design (35%), followed by multiple baseline/probe design (30%), alternating treatments design (5%), and combined design (5%).

Almost all included studies (95%) measured and visually presented the changes in the occurrence of challenging behavior across conditions (i.e., baseline and intervention). Replacement behavior was measured and reported in 16 studies included (80%), however, one quarter of these studies (n = 4) did not collect data on replacement behavior during the baseline condition. Therefore, behavior change in the use of appropriate communication response cannot be traced for these studies. Beyond targeting challenging behavior or replacement behavior as dependent variable, three studies also measured other child outcomes, such as positive facial expression, academic engagement, and “appropriately engaged” behavior.

Either live or video observation was used to measure challenging behavior in all included studies. Challenging behavior was most frequently measured and expressed as the percentage of intervals with challenging behavior (60%), which was calculated by taking the number of intervals with challenging behavior, dividing by the total number of intervals and multiplying by 100. Frequency counts of challenging behavior were used as outcome measures in four studies (20%), and rate of challenging behavior (e.g., number of challenging behaviors per minute) was used in the other four studies (20%).

Five studies (25%) reported the fidelity of implementation/treatment in terms of FCT agent’s behavior (e.g., who taught the communicative responses). Treatment fidelity was measured on the basis of either live or video observation of FCT agent’s behavior. Four studies named each behavior for which fidelity data were collected. Half of these studies (n = 2) reported fidelity data for each child participant, and four studies reported fidelity data across conditions.

Nine studies (45%) assessed generalization or maintenance of FCT effectiveness. However, social validity was only measured in three studies (15%). Teachers reported their satisfaction, acceptability, or feasibility of the intervention procedures in two studies. Naïve practitioners were asked to rate children’s behavior after watching several randomly selected baseline and intervention videotaped sessions to examine behavior changes in one study. Overall, findings from these studies indicated that FCT was feasible and acceptable to implement in early care and education settings, and was reported to be effective for decreasing challenging behavior.

What Works Clearinghouse Standards

In order to evaluate the strength of evidence, the 20 included studies were reviewed against the WWC Pilot Single-Case Design Standards (What Works Clearinghouse, 2017). No studies received a rating of Meets Design Standards without Reservations, eight studies received a rating of Meets Design Standards with Reservations, and 12 studies received a rating of Does Not Meet Design Standards. The main indicator that resulted in these findings was related to the number of data points per phase to qualify as an attempt to demonstrate an effect. Specifically, five studies did not provide at least three data points in a phase while meeting other design criteria. Other design issues identified included the use of non-experimental designs that did not provide three demonstrations of treatment effect (e.g., A-B and A-B-C designs), and low IOA (e.g., less than 20% of data points, or does not meet minimal IOA thresholds).

Following the application of WWC design standards, these eight studies that were rated as meeting design standards with reservations were subjected to visual analysis. For reversal/withdrawal design studies, child was the unit of analysis; whereas study served as the unit of analysis for multiple baseline design studies. One included study employing a multiple baseline design was rated as providing moderate evidence of treatment effect. One included studies employing a reversal/with-
drawal design with only one young child demonstrated strong evidence of treatment effect. Two reversal/withdrawal design studies with more than one young child provided strong evidence of effectiveness for only one child per study.

As noted previously, the WWC has encouraged single-case design researchers to identify evidence-based practices using the “5-3-20” threshold. Based on the findings from this review, the FCT research with young children exhibiting challenging behavior conducted in early care and education settings did not meet the evidence-based practice “5-3-20” threshold according to the WWC standards. Although one criterion was met (i.e., at least three different research teams with no overlapping authorship at three different institutions), only four single-case design studies met both design and evidence standards. Further, the number of young children was not sufficient to currently classify FCT as an evidence-based practice for the treatment of challenging behavior of young children with disabilities in early care and education settings based on the WWC standards.

**Magnitude of Treatment Effect**

We estimated effect sizes for the single-case design studies included in the review (no matter whether the study met the WWC standards or not), and a total of 13 studies were involved in effect size calculation. Effect sizes could not be calculated for seven studies, either due to the design (e.g., alternating treatment) or the intervention included variations and it was unclear which variation to focus the effect size on (e.g., presence or absence of picture cards; Hines & Simonsen, 2008). We found that FCT had a positive impact on the challenging behavior of young children. Overall, the mean PND was 74.09% (median = 75.00%; range = 25.46% to 100.00%) and the mean Tau-U was -.81 (median = -.81; range = -1.00 to -.52) across studies (see Table 3).

In order to provide a more accurate estimate of the overall effect, we conducted a random-effects meta-analysis with the studies that met both WWC design and evidence standards to ensure that the weighted mean effect size was robust with regards to methodological rigor and congruence with visual analysis (i.e., if visual analysis found no functional relationship, the effect size should not be calculated). Therefore, only four studies were included in the meta-analysis (Durand, 1992; Gibson et al., 2010; Radstaake et al., 2013; Volkert et al., 2009). Figure 3 illustrates the range of effect sizes and confidence intervals across these studies. Based on the results of meta-analysis, FCT interventions significantly decreased challenging behavior exhibiting by young children with disabilities when FCT was delivered in early care and education settings (weighted mean Tau-U = -.99; 95% CI −1.42 to −.56).

**Discussion**

Over the past three decades, FCT has been implemented and evaluated as an intervention designed to address challenging behavior with a range of populations, in a variety of settings, involving a myriad of professionals (Durand & Moskowitz, 2015). This synthesis was conducted to review the status and rigor of FCT research with young children exhibiting challenging behavior in early care and education settings. A secondary purpose was to explore the magnitude of experimental effects within and across studies. Specifically, this review sought to determine whether there was a sufficient evidence base to classify FCT as an evidence-based practice to reduce challenging behavior for young children in early care and education settings based on the WWC standards.

The majority of studies meeting review inclusion criteria provided general descriptions of the participants, settings, intervention characteristics, and primary outcomes/dependent variables. Less explicit descriptions were pro-
vided about the individuals who implemented the intervention (e.g., years of experience, training and certification, educational level), which compromises generalizability as it is unclear whether or not certain implementer characteristics moderate FCT effectiveness. Although concerning, this finding is consistent with a previous review examining early childhood implementation of, and implementer involvement in, FBA and function-based interventions (Wood, Drogan, & Janney, 2014).

Studies included in this review involved young children with a variety of disabilities and a range of challenging behavior. Although the definitions of challenging behavior varied across studies, young children targeted for FCT were more likely to exhibit externalizing behaviors that disrupt regular classroom activities, including peer and adult aggression and tantruming. This is consistent with previous findings that young children with externalizing versus internalizing behaviors are more likely to be referred for intensive and individualized intervention, such as FCT (McKenna, Flower, Kim, Ciullo, & Harring, 2015).

Although researchers primarily conducted the FBA components in the reviewed studies, early childhood classroom teachers were actively involved in the process of identifying the function of children’s challenging behavior. Further, they played a primary role in teaching children the identified replacement behavior. This finding supports a previous review of practitioner-implemented FCT indicating that teachers can successfully implement FCT in schools (Andzik et al., 2016). Furthermore, instruction in replacement behavior occurred in regular classrooms for more than half of the young children across the included studies. This encouraging finding suggests that many young children are receiving FCT in natural settings and classroom teachers are actively participating in FCT interventions. Teacher involvement in FBA and teaching replacement behavior is essential for improving the implementation and acceptability of the FCT intervention (Wood et al., 2014).

The purpose of FCT is to replace challenging behavior with appropriate forms of communicative response that serve the same function (Durand & Moskowitz, 2015). Almost all studies included in this review attempted to visually demonstrate a reduction in challenging behavior. However, behavior change in the use of alternative communication responses could not be ascertained in several studies due to the absence of baseline data or no data collected on alternative communication responses. Therefore, the extent to which FCT increases the use of alternative communication responses cannot be determined for these studies. A decrease in the occurrence of challenging behavior cannot automatically imply an increase in the use of alternative communication responses. Future research should include measures of alternative communication responses across conditions to examine the effect of FCT on alternative communication responses.

Although the majority of studies included in the review were published after 2000, few studies provided information on treatment fidelity. Measuring treatment fidelity strengthens the internal and external validity of the study by demonstrating that the intervention was implemented as intended and therefore is essential to making valid conclusions about the study (Ledford & Wolery, 2013). The limited reporting of treatment fidelity significantly limits the generalization and findings and effect size estimates of the studies included would likely be more robust if the FCT was implemented as intended.

Compared with the number of included studies reporting treatment fidelity, even fewer studies reported a social validity measure. However, the number of studies reporting generalization or maintenance data is larger than what was found in an earlier review of FCT studies for children with ASD (i.e., Mancil, 2006). Almost all included studies reported adequate levels of IOA, however, two studies reported that IOA data were collected for fewer than 20% of sessions and IOA did not meet minimal thresholds specified in the WWC standards in two studies. It is worth noting that more than half of the included studies did not document whether IOA data were collected in each condition (e.g., baseline, intervention).

Across all single-case design studies included in our synthesis, not one met the WWC design standards without reservations. The
most common design limitation was the inadequate number of data points collected for each phase. For example, reversal/withdrawal or multiple baseline design studies with fewer than three data points in any phase would automatically result in receiving a rating of Does Not Meet Design Standards. Inadequate number of data points per phase was also the main reason studies did not meet design standards in a previous review applying the WWC standards for evaluating single-case research related to interventions for improving daily living skills of individuals with ASD (Hong et al., 2015). These findings suggest the issue of inadequate number of data points per phase might not be unique to FCT single-case design research.

Across the eight studies that met the WWC design standards with reservations, four were rated as providing either strong or moderate evidence of treatment effect. Intervention effectiveness was pooled for each study that met both WWC design and evidence standards. Overall, a large weighted mean effect size was found, suggesting that FCT interventions were effective in reducing the challenging behavior of young children with disabilities in early care and education settings. However, given the small number of FCT studies that met WWC design and evidence standards, only nine young children were involved in these studies, which did not meet the required number of cases (i.e., \( \geq 20 \)) in the WWC’s “5-3-20” threshold. Therefore, FCT cannot currently be characterized as an evidence-based practice for the treatment of challenging behavior of young children with disabilities in early care and education settings, according to the WWC standards.

**Limitations of the Synthesis**

The current study adds to existing literature by quantitatively synthesizing FCT research conducted in early care and education settings for young children with challenging behavior. However, several limitations should be noted. First, only articles published in a peer-reviewed journal were included in this synthesis, which might lead to publication bias or the tendency to highlight only positive results (Easterbrook, Gopalan, Berlin, & Matthews, 1991). Second, this synthesis only included FCT research conducted in early care and education settings, therefore findings may not be generalizable to similar studies conducted in other settings, such as home or clinic settings. Third, the strict WWC’s “5-3-20” threshold prevents FCT from being currently considered an evidence-based practice because fewer than 20 young children were involved in this body of research, and fewer than five studies included met both WWC design and evidence standards. Fourth, consensus has not been reached on the optimal methods for computing effect sizes and using meta-analytic techniques for single-case design studies, therefore the effect size estimates and the results of this meta-analysis should be interpreted with caution.

**Implications for Future Research**

More studies with methodological rigor are needed to verify that FCT interventions conducted in early care and education settings are effective in decreasing challenging behavior and increasing communicative response behaviors. Efforts should be made to improve the rigor and quality of the single-case experimental design procedures used, including collecting more data points in each phase to obtain a stable data pattern, measurement of treatment fidelity, evaluation of social validity, and the provision of at least three attempts to demonstrate an intervention effect. Considering the lack of replacement behavior measures, future research should also collect data on replacement behavior, in order to examine the effect of FCT on the use of replacement behavior. Furthermore, replacement behavior should be measured across conditions and phases to enable the calculation of effect sizes.

This review and meta-analysis examined whether and to what extent FCT was effective in decreasing challenging behavior of young children with disabilities when FCT was delivered in early care and education settings. Results are promising, but future research is needed to: (a) confirm FCT as an evidence-based practice, and (b) further explore whether intervention effectiveness is related to participant characteristics (e.g., categories of disability, age, and language skills), intervention features (e.g., intervention agent, fi-
delity of intervention implementation, dosage of treatment), and intervention outcomes.

References
(References marked with an asterisk indicate studies included in the systematic review)


Gage, N. A., Lewis, T. J., & Stichter, J. P. (2012). Functional behavioral assessment-based interventions for students with or at risk for emotional and/or behavioral disorders in school: A hierar-


R Development Core Team (2012). *R: A language


Received: 25 July 2018
Initial Acceptance 20 September 2018; Final Acceptance 15 December 2018
Increasing Peer-Directed Vocal Manding of Young Children with Autism Spectrum Disorder

Kayla Centone
Temple University and Elwyn, Inc.

Matt Tincani
Temple University

Abigail Dilks
Elwyn, Inc.

Abstract: Manding for preferred items and activities is a critical skill for young children with autism spectrum disorder (ASD). However, little research has evaluated procedures to teach children with ASD to mand with their peers. This study evaluated the effects of a teaching package involving least-to-most prompting to increase peer-directed vocal manding using carrier phrases (e.g., “I want . . .”) of three preschool-aged children with ASD in a classroom setting. Following a preference assessment and baseline, experimenters taught children to vocally request toys from their peers through a combination of least-to-most prompting and differential reinforcement of peer-directed manding. Results indicated that each child acquired independent, peer-directed vocal manding with carrier phrases; manding responses generalized across different peers; and manding responses maintained in the classroom setting up to four weeks after the study concluded.

Many young children with autism spectrum disorder (ASD) have delayed or absent speech (American Psychiatric Association, 2013; Shriberg, Tomblin, & McSweeny, 1999). Young children’s inability to express their wants, needs, and preferences through speech can increase problematic behaviors, including aggression, crying, and self-injury (Battaglia, 2017; Carr & Durand, 1985). Absence of social and communicative skills may also reduce children’s participation in peer social networks (Chamberlain, Kasari, & Rotheram-Fuller, 2007) and limit their opportunities to learn new skills (Sundberg & Partington, 1998). Thus, strategies to promote communication, including speech, are an essential component of evidence-based intervention programs for young children with ASD (Wong et al., 2015).

Skinner (1957) defined the mand as “a verbal operant in which the response is reinforced by a characteristic consequence and is therefore under the functional control of relevant conditions of deprivation or aversive stimulation.” (Skinner, 1957; p. 35). More colloquially, mands are requests that take a variety of forms, such as asking for a preferred item, saying no to a non-preferred activity, making a choice, or asking for a break from a demanding task. The “conditions of deprivation and aversive stimulation” described by Skinner are also known as motivating operations (MOs) (Laraway, Snyderski, Michael, & Poling, 2003). MOs may momentarily increase the value of specific reinforcers and increase the likelihood of behaviors that have produced them in the past. For example, being deprived of a preferred toy (MO) momentarily increases the reinforcing value of the toy for a child with ASD, along with the likelihood of behaviors that have produced the toy in the past, (e.g., asking for the toy, grabbing the toy).

While manding is an important communication skill for children with ASD to attain their basic wants and needs, the social deficits of ASD may prevent children from learning to mand from their natural environment without explicit instruction (Tincani, Bondy, & Crozier, 2011). The majority of studies published
to date have focused on teaching children with ASD to engage in adult-directed manding (e.g., Betz, Higbee, & Pollard, 2010; Charlop-Christy, Carpenter, Le, Leblanc, & Kellet, 2002; Hernandez, Hanley, Ingvarsson & Tiger, 2007; Kelley, Shillingsburg, Castro, & Addison, 2007; Plavnick & Vitale, 2014; Shafer, 1994). This focus on teaching adult-directed manding is logical given that most children naturally mand to adults more frequently than they mand to their peers, which may be explained by the more immediate, contingent reinforcement of manding provided by adults (Paden, Kodak, Fisher, Gawley-Bullington, & Bouxsein, 2012). In contrast, when manding to a peer, there is likely to be a longer delay to reinforcement, or no reinforcement at all. However, directing mands to peers is also a critical functional communication skill for several key reasons (Kodak, Paden, & Dickes, 2012). Manding to peers creates opportunities for peer socialization and interaction, which is typically lacking in children with ASD (Chamberlain et al., 2007). Since peers are frequently together in classrooms and community settings, these are natural opportunities for the children to engage in communication, social interaction and play, obviating the need for teachers to contrive communication opportunities through structured lessons. Furthermore, having functional verbal behavior such as manding to peers potentially decreases problematic behaviors that may occur surrounding sharing and turn taking (Durand & Carr, 1992).

Despite the importance of peer-directed manding in functional communication of children with ASD, few studies have sought to explicitly teach peer-directed manding. Paden et al. (2012) taught children with ASD who used the Picture Exchange Communication System (PECS; Frost & Bondy, 2002) to engage in peer-directed mands with PECS. Within a reversal design, these experimenters taught the two children to exchange a card with their peer for access to a selected toy. During baseline, participants were permitted to mand for a preferred item to an adult or a peer. During intervention, all adult-directed mands were placed on extinction while peer-directed mands were reinforced using differential reinforcement of alternative behavior (DRA). A system of least-to-most prompting was used, as necessary, to prompt peer-directed manding during the intervention. During baseline, participants displayed only adult-directed manding. During intervention, peer-directed manding increased to up to four times per minute, while adult-directed manding substantially decreased. Brief social interactions with peers (e.g., playing with a toy together) also increased during intervention, indicating an overall increase in peer interaction as an additional benefit of the intervention.

In a similar experiment, Kodak et al. (2012) used a reversal within a multiple baseline design to teach peer-directed manding to two children with ASD who used PECS for communication with adults. During baseline, experimenters reinforced participants’ manding towards adults and peers. During intervention, adult-directed manding was placed on extinction and prompting and reinforcement was used to increase peer-directed manding. The participants showed no peer-directed manding during baseline. During intervention, independent peer-directed manding increased while adult-directed manding decreased. The study successfully replicated Paden et al. (2012), providing further support for DRA plus extinction and prompting in promoting peer-directed manding. Additionally, these authors evaluated treatment effects on manding to a novel peer, manding by traveling across the classroom, and manding to a novel peer in a more naturalistic setting. Both participants were able to engage in peer-directed manding from distances across the classroom and with novel peers in a naturalistic setting. This finding supports naturalistic teaching strategies to promote generalization of peer-directed manding when implemented in applied settings.

Hanney (2015) conducted a two-part study comprised of conditioning peers as reinforcing in the first study and teaching peer-directed manding in the second study. In the second study, the experimenter taught four participants with ASD to engage in vocal manding to novel peers and previously conditioned peers who had been paired with the participants’ preferred stimuli. The experimenters used most-to-least prompting with a time delay to teach peer-directed manding, increasing the time delay prior to prompt de-
livery, and systematically fading out prompts to facilitate independent mands. None of the participants engaged in peer-directed manding during baseline. During intervention, three out of four participants displayed high levels of vocal manding to both the novel and conditioned peers.

There are several empirically-supported treatments to teach communication, including manding, to children with ASD (Wong et al., 2015). However, despite the benefits of peer-directed communication, the majority of research to date has focused on empirically validating strategies for increasing manding to adults rather than manding to peers. Three studies have shown that children with ASD can be taught to mand to their peers through a combination of behavioral teaching procedures, including differential reinforcement and prompting (Hanney, 2015; Kodak et al., 2012; Paden et al., 2004); however, more research is needed to confirm whether these strategies are effective in teaching peer-directed manding. Moreover, only one of these studies taught vocal manding towards peers (Hanney, 2015). Since vocal communication is the most natural and desirable form of verbal behavior, more studies on teaching vocal manding are needed. Therefore, in this study we sought to evaluate the effects of a teaching package with differential reinforcement and least-to-most prompting on peer-directed vocal manding with carrier phrases (e.g., “I want . . .”) of children with ASD. Manding with carrier phrases was selected as the target skill because in their classroom setting, all students, including the study’s participants, were encouraged to mand with carrier phrases, and thus participants were taught to use carrier phrases to increase the likelihood of generalization in the classroom setting. Unlike most previous research on this topic, our study explicitly examined procedures for establishing peer-directed manding using vocal speech, as opposed to an augmentative communication modality (e.g., PECS; Frost & Bondy, 2002). Specifically, we sought to answer the following research questions:

1. What are the effects of a teaching package with differential reinforcement and least-to-most prompting on increasing independent, peer-directed vocal manding of three young children with ASD?
2. To what extent, if any, do gains in independent, peer-directed vocal manding maintain after withdrawal of intervention?

Method

Participants

Three preschool-aged children with ASD, Jason, Tommy, and Noah, served as participants in the study. Four of their classroom peers with ASD served as peers to be the recipients of peer-directed manding in the study. The researchers selected the peers based on a classroom observation checklist, which evaluated their ability to independently follow simple directions (e.g., “Sit down,” “Hand me the toy,” or “Give the toy to . . .”). The peers did not exhibit significant challenging behavior and were able to engage in parallel or cooperative play for at least 5-min. To program for generalization, a random number generator was used to randomly match the peers with different participants during each baseline and intervention session.

Through direct observation by the first author and staff reports, Jason, Tommy, and Noah showed some vocal communication and the ability to independently and spontaneously vocally mand for a variety of tangibles with adults in their natural environment; however, they showed little vocal manding with peers. Like their peers, each participant demonstrated the ability to follow simple, one-step receptive instructions. The Adaptive Behavior Assessment System-second Edition (ABAS-II; Harrison & Oakland, 2003) and the Childhood Autism Rating Scale, Second Edition (CARS-2; Scholper, Van Bourgondien, Wellman, & Love, 2010) were used to assess the participants’ adaptive behavior and communication skills prior to baseline. The participants’ classroom teacher was the rater for the CARS-2. The participants’ parents were raters for the ABAS-II.

Jason was a 4-year-old boy diagnosed with ASD and attention deficit / hyperactivity disorder – combined type, who exhibited a language impairment and required substantial communicative support. Jason received a scaled score of 3 on the general adaptive com-
posite (GAC) scale of the ABAS-II, in the extremely low range for communication. He fell in the mild-to-moderate range of symptoms of ASD when assessed using the CARS-2, with a raw score of 30.

Tommy was a 5-year-old boy diagnosed with mild ASD who also required substantial communicative support. He scored in the extremely low range of the GAC scale of the ABAS-II with a scaled score of 4. He fell into the mild-to-moderate range of symptoms of ASD according to the CARS-2 with a raw score of 36.

Finally, Noah was a 5-year-old boy diagnosed with mild ASD requiring support in the area of social communication. He scored in the extremely low range of the GAC scale of the ABAS-II with a scaled score of 5. He fell into the mild-to-moderate range of symptoms of ASD according to the CARS-2 with a raw score of 31.

Settings and Materials

All of the study sessions occurred at a therapeutic after-school program for children with ASD located in the suburbs of a large city in the mid-Atlantic region of the US. A nonprofit organization that serves children and adults with a variety of disabilities ran the program. The program placed students in instructional groups of three to four children, with two staff per group. All students in each group had ASD. The daily schedule was comprised of lunch, facilitated choice play, story time, arts and crafts, group games, sensory activities, playground time, and other cooperative play stations. The children’s treatment plan goals included cooperative or parallel play, requesting from adults and peers, initiating social interactions, responding to social interactions, and independently transitioning between activities.

The study occurred in two settings within the program, the children’s classroom and a separate room. The classroom had three windows and one door entrance. There were two tables in the room, one used for children activities and the other used for staff paperwork and materials. The room also had one wall of cubbies and two small shelving stands. There was artwork, a white board, bulletin board, schedules and visual pictures on the walls. The separate room was a small room that was completely empty. It had two windows, one door entrance, and a closet, which was locked any time the participants were in the classroom.

Baseline sessions always occurred in the separate room. To program for generalization, intervention sessions occurred in both the classroom with all peers in the room and in the separate room with only the participant, one peer, the experimenter, and an additional data collector present during interobserver agreement (IOA) and procedural fidelity checks. The two settings used for intervention sessions were randomly alternated using a random number generator. Maintenance probes only occurred in the classroom setting. Of the 20 intervention sessions for Jason, 11 sessions occurred in the classroom and nine occurred in the separate room. Of the 14 sessions for Tommy, eight occurred in the classroom and six occurred in a separate room. Of the eight sessions with Noah, three occurred in the classroom and five occurred in the separate room.

The materials for the study were toys determined from each participant’s preference assessment that were used to reinforce adult- and peer-directed mands. Procedures for the preference assessment are described below.

Dependent Variable

The dependent variable was the participants’ percentage of trials with independent, vocal peer-directed mands with carrier phrases. An independent peer-directed mand was recorded if the participant spontaneously used a sentence frame to emit the target mand, directing his eye gaze toward the peer, without any prompting from the experimenter. Generalized sentence frames such as, “Can I have the (item name),” “I want (item name),” or “I need (item name),” were recorded as independent mands as long as the participant emitted at least three words, including a pronoun, verb, and the name of the target item. The participant was required to say the item name for the response to be recorded as an independent peer-directed mand. For example, “I want that,” or “My turn,” were not recorded as independent peer-directed mands. A prompted peer-directed mand was recorded if
the participant correctly said the target mand or a fragment of the target mand to a peer only after a gestural or vocal prompt from an adult (e.g., “I want . . .”). A nonresponse was recorded if the participant engaged in no vocal response or if the participant made any other response that did not match what was prompted. There were no nonresponses observed during the study.

Data Collection

The experimenter, the first author, was the primary data collector throughout the study. She was a master’s student in applied behavior analysis who worked as a clinical case manager at the program and lead classroom socialization groups for children in the program, including those who participated in the study. The data sheet was divided into a table of 10 trials. Each row of the table represented one trial. The data collector would circle either, “independent,” “prompted,” or “nonresponse,” for each trial. The experimenter converted the total number of independent peer-directed vocal mands into a percentage by dividing the number of independent mands by the total number of opportunities for the session (ten trials) and then multiplying by 100.

Interobserver Agreement

Two additional data collectors collected data for IOA. The first data collector was the after-school program’s clinical supervisor who was a Board Certified Behavior Analyst with 10 years of experience working with children with ASD. The second data collector was a lead-staff member who was in graduate school for clinical counseling and had three years of experience working with children with ASD. Prior to the start of the study, both data collectors reviewed the operational definitions and were trained using practice sessions with children who attended the program. They were taught to record data on independent mands, prompted mands and nonresponses until they reached 100% agreement in practice sessions. Data collectors recorded IOA data during 33% of baseline sessions for each participant. IOA data were collected during 30% of Jason’s intervention phase, 29% of Tommy’s intervention phase, and 25% of Noah’s intervention phase. The experimenter calculated IOA by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. IOA was 96.6% (range, 90–100%) during baseline phases and 95% (range, 80–100%) during intervention phases.

Experimental Design

This study used a single-case, multiple-probe across participants design (Cooper, Heron, & Heward, 2007). Each participant was exposed to a continuous baseline for three sessions immediately prior to the start of the intervention phase. The criterion for beginning intervention was a minimum of three data points showing low, stable levels of responding (Horner & Baer, 1978). When the first participant, Jason, showed low, stable responding, and no improvement in his baseline phase, the intervention was introduced. Once the first participant reached at least one data point of 90% independent mands, the intervention was introduced to the second participant, Tommy, whose baseline also showed low, stable responding. Once Tommy reached at least one data point of 90% independent mands, the intervention was introduced to the third participant, Noah, whose baseline also showed low, stable responding.

Procedure

Preference assessment. Experimenter classroom observations prior to the study determined that all three participants had restricted interests focusing on toy trains and cars. Before baseline, the experimenter administered a one-time, seven-item multiple stimulus without replacement (MSWO) assessment (Leon & Iwata, 1996) comprised of toy trains and cars to determine items for peer-directed mand training. Additionally, the experimenter conducted a probe to ensure that each participant could say the name of each preferred item (i.e., “What is this?”). For each participant, the experimenter identified one most highly preferred item by the assessment. Jason’s most highly preferred item was a Mickey Mouse toy car, Tommy’s was a Thomas the Train toy, and Noah’s was a Percy train toy. Following the preference assessment, all
three preferred items were placed into a clear plastic bag, which was kept in an office and not accessible to the participants during any other time of the day besides the study sessions.

**Baseline.** The purpose of baseline was to evaluate the extent to which participants could perform peer-directed vocal manding with carrier phrases in the absence of any explicit teaching procedures. Each baseline session was comprised of 10 trials, which lasted approximately 30-s each, for a total session duration of 5-min. During each baseline session, the target participant and a peer were in the separate session room, along with the experimenter and the secondary observer. The four peers were randomly rotated across each baseline session. Before beginning each session, the experimenter presented the most highly preferred item to the participant or, if the participant vocally indicated that he wanted a different item from the clear plastic bag, the experimenter allowed the participant to choose another item from the bag. The experimenter provided this choice to ensure that the participant was motivated to request during the session. After the participant selected a toy, the experimenter placed the remaining toys in the plastic bag, which was removed from the room so the toys were no longer visible during the session.

The experimenter seated the target participant and peer facing each other on the floor approximately 1–2 feet apart. The experimenter gave the target participant's chosen toy to the peer upon the start of the session and said to the peer, “You can play with this.” The experimenter then told the participant, “If you want the (item name), you need to say, I want (item name).” There were no additional prompts given to the participant for peer-directed manding with carrier phrases. If the peer tried to give the toy to the participant before the participant manded, the experimenter blocked the exchange and reminded the peer that it was time for him to play with the toy. Contingent on an independent vocal peer-directed mand, the experimenter prompted the peer to provide the target participant with the toy for 30-s. If the participant performed an unprompted peer-directed vocal mand, the experimenter prompted the peer to provide the target participant with the toy. The experimenter also provided behavior specific verbal praise (e.g., “Good asking!”) contingent on a correct mand only during the first trial of each session to avoid verbal praise from adults functioning as conditioned reinforcement for mands.

If independent manding did not occur within 5-s, the experimenter prompted the participant to mand for the toy using least-to-most prompting, comprised of a gestural prompt (i.e., pointing to the peer), a partial verbal prompt (i.e., “I . . .”), and a full verbal prompt (i.e., “I want the (item name).”), with 5-s in that trial, until the next opportunity for the participant to mand. Any mands for the item directed to the experimenter were ignored during baseline, in that the participant was not given the item and the experimenter withheld any prompts or attention.

**Mand Training.** The purpose of mand training was to increase participants’ peer-directed vocal manding with carrier phrases through differential reinforcement and least-to-most prompting. As in baseline, the target participant and peer sat facing each other on the floor approximately 1–2 feet apart for 10 trials, with the peer holding the participant’s selected toy at the beginning of each trial. At the beginning of the first trial, the experimenter told the participant, “If you want the (item name), you need to say, I want (item name).” However, during mand training, the experimenter only gave this instruction during the first trial of all sessions until the participant reached at least 50% independence in peer-directed manding during a session, after which the instruction was no longer provided. Following the instruction, the experimenter gave the participant access to the selected toy, non-contingently, for 30-s. Then, the experimenter prompted the participant to give the toy to the peer by manually guiding the removal of the toy and handing it to the peer and providing verbal prompting (e.g., “It’s [peer’s name’s] turn”), as needed. No additional instructions were given to the peer. At the end of 30-s peer access, the experimenter gave the participant 5-s to independently vocally mand for the toy to the peer. As in baseline, if the participant performed an unprompted peer-directed vocal mand, the experimenter prompted the peer to provide the target participant with the toy. The experimenter also provided behavior specific verbal praise (e.g., “Good asking!”) contingent on a correct mand only during the first trial of each session to avoid verbal praise from adults functioning as conditioned reinforcement for mands.
in between each prompt. If the participant still did not respond after one full verbal prompt, a second full verbal prompt was given. If the participant did not respond after a second full verbal prompt, the peer was allowed to play with the toy for the remainder of the 30-s trial, and then a new trial began. After two trials of the participant not responding after two full verbal prompts, the experimenter asked the peer to select a new item from the bag of toys. If after two trials of the participant not responding after two full verbal prompts with the toy, the experimenter presented the third toy. If there was no responding after two full verbal prompts within two trials with this toy, the session would be terminated early; however, participants always responded prior to two full verbal prompts with the third toy, and therefore sessions were never terminated early.

**Maintenance probes.** The experimenter conducted one maintenance probe 2 weeks after the end of intervention and another two maintenance probes 4 weeks after the end of intervention. These probes evaluated for maintenance of the peer-directed manding behavior over time. Maintenance probes were identical to baseline procedures, except all probe sessions were conducted in the classroom setting.

**Procedural Fidelity**

The same data collectors who collected IOA data on the dependent variable also assessed procedural fidelity using a checklist of the baseline and intervention procedures. The checklist for baseline contained six items, whereas the checklist for the intervention contained eight items. The experimenter trained data collectors to collect procedural fidelity data by reviewing the task list steps and collecting data during practice sessions with children who were not included in the study. Data collectors reached 90% IOA during practice sessions prior to collecting procedural fidelity data in baseline. Procedural fidelity data were collected during 33% of baseline sessions for all participants, and during 25% of Jason’s intervention phase, 29% of Tommy’s intervention phase, and 33% of Noah’s intervention phase. Procedural fidelity across subjects averaged 94.3% (range, 83–100%) during baseline and 89% (range, 62.5–100%) during intervention.

The steps of the intervention checklist most frequently missed were the participant and peer seated 1–2 feet apart (step 1), which was implemented correctly for 41.7% of opportunities, and the experimenter correctly implementing the prompt hierarchy (i.e., beginning with a gestural prompt rather than a verbal prompt) (step 6), which was implemented correctly for 81.8% of opportunities. The participants not sitting 1–2 feet apart, step 1, often occurred in the separate classroom where the participants had more room to move around, and thus they were often positioned farther than 1–2 feet away from each other when manding. Excluding step 1, average procedural fidelity during intervention was 96.4% (range, 71.4–100%). Additionally, staff exhibited more procedural errors during initial intervention sessions with Tommy, whose procedural fidelity averaged 80% (range, 63–100%) than they did with Jason and Noah. Both Jason’s and Noah’s procedural fidelity averaged 95.2% (range, 88–100%).

**Social Validity**

We administered two social validity questionnaires, one to staff who worked with the participants and another to the participants and peers. We developed these social validity questionnaires based on Schwartz and Baer’s (1991) conceptual framework to discover whether the goals, procedures, and outcomes of the intervention were important to primary and secondary consumers.

**Staff questionnaire.** The staff questionnaire was administered to two staff persons who worked with the participants and peers at the therapeutic after-school program, who were asked to respond by giving a Likert scale rating (1 = strongly disagree; 5 = strongly agree) to the following questions. (a) Peer-directed mands are an important skill for preschool children with ASD to learn; (b) I liked the procedures used in this study to teach peer-directed mands; (c) I thought the procedures used in the study were immediately effective; and (d) I think the procedures used in this study are still showing effectiveness following a few weeks since the study has ended.
Participant and peer questionnaire. The participant and peer questionnaires were comprised of two questions, which asked if they (a) liked learning to ask to share with his friends and if (b) they thought sharing is being a good friend. The participants and peers answered each question by pointing to a happy face, neutral face or sad face.

Analytic Approach

We conducted the following analyses of the data. First, we graphed and visually analyzed data across phases. Specifically, following collection of sufficiently stable baseline data, we visually determined how changes in level, trend, and variability corresponded with application of intervention, along with immediacy of experimental effect, overlap of data between phases, and consistency of data patterns between participants (Kratotchwill et al., 2013). Given presence of experimental control confirmed through visual analysis, we generated effect size estimates with Tau - U, a non-parametric technique based on non-overlapping data points that yields effect size estimates with confidence intervals (Parker, Vannest, Davis, & Sauber, 2011) using an online calculator (http://singlecaseresearch.org; Vannest, Parker, & Gonen, 2011). Using the baseline trend evaluation tool available in the online calculator, we determined that baseline corrections for trend were not necessary. We conducted a series of three A – B contrasts for each participant, where A was baseline and B was intervention and maintenance, along with calculating an omnibus effect size estimate for all three participants. We also calculated Hedge’s g to determine differences between the means of baseline and intervention and maintenance phases. Hedge’s g is an effect size estimate similar to Cohen’s d that corrects for bias within small sample sizes, such as with single-case designs (Levy et al., 2017).

Results

Figure 1 shows the percentage of independent peer-directed manding with carrier phrases during baseline, intervention, and maintenance phases. Data collected in the separate room are depicted with the diamond-shaped data markers, whereas data collected in the classroom are depicted with the circular data markers.

As shown in Figure 1, all participants displayed low levels of independent peer-directed manding during baseline. Implementation of the intervention corresponded with immediate increases in responding for all three participants. The initial slope in each participant’s intervention data suggests that it took six to seven sessions for each participant to fully acquire independent manding. Visual analysis indicates little or no overlap between conditions, with marked increases in levels of independent, peer directed manding with intervention. Visual inspection of data collected in the separate room (diamond shaped markers) versus data collected in the classroom (circular markers) suggests little overall difference between responding across these two locations. Maintenance data show that Jason’s and Noah’s higher levels of independent manding during intervention maintained in the classroom without any explicit instruction four weeks after intervention. Tommy’s level of independent manding was somewhat lower during maintenance than during the latter intervention sessions; however, they still maintained above baseline levels. An omnibus Tau - U coefficient was generated to compare baseline responding with intervention and maintenance responding. Tau – U across participants was .996 (95% CI [.651, 1.00], p < .0001). Hedge’s g was also calculated to compare baseline responding with intervention and maintenance responding; g = 3.87. Collectively, visual and quantitative analyses reflect strong experimental control and large treatment effects.

Jason. As displayed on the top panel of the graph in Figure 1, during baseline, Jason displayed a mean of 3% (range, 0–10%) independent peer-directed mands. Anecdotally, his verbal behavior in an attempt to gain the toy during baseline mainly consisted of crying, whining, grabbing towards the item, saying, “Hey,” or only stating the name of the item using one word. Once the intervention was implemented, Jason’s independent mands immediately increased to a mean of 82.5% (range, 30–100%). Jason reached the criterion of 90% independent manding in the sixth session. After reaching the criterion of 90% independence, Jason’s average indepen-
Figure 1. Percentage of peer-directed mands during baseline, intervention, and maintenance for Jason, Tommy, and Noah. The diamond-shaped data markers represent data collected in the separate room; the circular data markers represent data collected in the classroom.
dent peer-directed manding for sessions 6 – 13 was 94% (range, 80–100%). During maintenance, Jason’s manding continued at mastery levels, as his mean independence was 93.3% (range, 90–100%).

**Tommy.** As displayed in the middle panel of the graph in Figure 1, during baseline, Tommy displayed a mean of 21.6% (range, 10–40%) independent peer-directed mands. Anecdotally, Tommy’s verbal behavior in an attempt to gain the toy during baseline consisted of crying, whining, grabbing towards the item, saying, “Mine,” “My turn,” and, “Pass it to me.” Once the intervention began, Tommy’s independent mands increased to a mean of 78.5% (range, 40–100%). Tommy reached the criterion of 90% independent mands during the seventh intervention session. After reaching the mastery criterion, Tommy’s average independent peer-directed mands for sessions 7 – 14 was 90% (range, 80–100%). During maintenance probes, Tommy did not maintain his manding at mastery levels. His independent manding across three maintenance probes averaged 66.6% (range, 60–70%).

**Noah.** As displayed in the bottom panel of the graph in Figure 2, during baseline, Noah displayed a mean of 7.7% (range, 0–30%) independent peer-directed mands. Anecdotally, his attempts to gain the toy included grabbing for the item, screaming, crying, and stating the name of the item using one word. During the intervention phases, Noah’s independent mands immediately increased to a mean of 77.5% (range, 60–90%). Noah reached the mastery criterion of 90% independence on the sixth session. Noah remained at 90% independence for all sessions following the initial session of meeting the criteria. During maintenance probes, Noah maintained mastery criteria of 90% independence during the first probe, 2 weeks following the last intervention session. However, during the four week maintenance probe, Noah displayed 80% independence. His overall mean for maintenance probes was 83.3% (range, 80–90%).

**Staff questionnaire.** We asked two staff persons to provide a Likert scale rating (1 = strongly disagree; 5 = strongly agree) to five survey questions regarding the goals, procedures, and outcomes of the intervention. On average, the two staff persons gave a rating of 4.8 to all of the survey questions, suggesting that they held strongly positive perceptions of the intervention. Staff gave responses of 5 to all of the survey questions, except for the third question, “I thought the procedures used in the study were immediately effective,” to which each gave a response of 4.

**Participant and peer questionnaire.** The participant and peer survey was comprised of two questions, which asked each participant and peer if he (a) liked learning to ask to share with his friends, and (b) thought sharing is being a good friend. All peers and participants selected the “happy face” for both questions on the surveys. Anecdotally, the participants and peers seemed to enjoy participating in the study. They were often smiling and stated that they, “Wanted to play,” when it was time to begin a session.

**Discussion**

The purpose of the study was to evaluate the effects of a teaching package with differential reinforcement and least-to-most prompting on peer-directed vocal manding with carrier phrases of children with ASD. Results of the study support and extend previous research (Hanney, 2015; Kodak et al., 2012; Paden et al., 2012) in showing that a combination of these procedures increased the frequency of children’s independent, peer-directed manding compared to baseline, and that children were able to generalize their mands across different peers in the natural classroom setting. Furthermore, increases in children’s peer-directed manding maintained up to four weeks after the withdrawal of intervention. Finally, classroom staff who were surveyed rated the goals, procedures, and outcomes of the intervention favorably.

The first research question addressed the effects of a teaching package with least-to-most prompting on increasing independent, peer-directed vocal manding of three young children with ASD. Visual inspection of Figure 1 indicates that introduction of the intervention coincided with increases in independent, peer-directed vocal manding across the three participants at three different points in time, suggesting a functional relationship between the intervention and increased responding. A
learning curve is evident, where it took nine, eight, and six sessions for Jason, Tommy, and Noah, respectively, to reach 100% independent responding. As shown in Figure 1, Tommy’s overall level of independent manding was somewhat lower than Jason’s and Noah’s; however, his level of independent responding may have been affected by lower levels of procedural fidelity during his initial intervention sessions, as described in the results.

The second research question addressed the extent to which independent, peer-directed vocal manding responses maintained after withdrawal of intervention. As shown on the right side of Figure 1, for Jason, peer directed vocal manding maintained at intervention levels for 2 and 4 weeks after intervention was withdrawn. A similar pattern of responding during maintenance was observed for Noah. Although Tommy’s level of independent responding during maintenance was higher than in baseline, his peer directed manding decreased somewhat following withdrawal of intervention for reasons that are unclear.

Limitations

The procedures and results of the study suggest a possible boundary of the intervention for practitioners who wish to teach peer-directed manding to young children with ASD. As with participants in the Paden et al. (2012) and Kodak et al. (2012) studies, participants in the current study displayed adult-directed manding prior to intervention. Therefore, it was unnecessary for the experimenters to employ intensive teaching procedures to establish basic manding repertoires during the intervention (e.g., Frost & Bondy, 2002). Rather, a teaching package with least-to-most prompting was sufficient to transfer existing adult-directed manding to peers. However, children with ASD who lack strong adult-directed manding skills prior to intervention may benefit from more intensive teaching procedures, including most-to-least prompting and multiple instructors, to establish robust peer-directed manding repertoires.

We did not measure participants’ social interactions besides peer-directed manding in the study, nor did we measure peers’ social responses in the study. Consequently, we cannot report whether the intervention positively affected other peer-to-peer social interactions, such as greeting, commenting on activity, questioning, or responding to an inquiry (Owen-DeSchryver, Carr, Cale, & Blakeley-Smith, 2008). Since participants and peers were required to take turns with and share each toy during the intervention, and these skills were only weakly present during baseline, it is likely that the intervention facilitated turn taking and sharing in the context of peer-directed manding. However, we acknowledge that these and other critical peer-to-peer social interactions are likely to increase only in the context of a comprehensive intervention program that extends well beyond mand training (Barton & Harn, 2012). We also did not measure challenging behavior in the study; however, anecdotally, during baseline the participants engaged in a variety of challenging behaviors in order to access the toys (e.g., crying, whining), and these behaviors appeared to diminish during intervention as participants acquired peer-directed manding.

Since the preferred items in the study were always in the participants’ eyesight, the items themselves may have exercised a degree of stimulus control over the manding response (Bondy, Tincani, & Frost, 2004). In contrast, naturalistic classroom situations often involve children manding items that are not physically in view. Furthermore, each peer delivered the toy to the participant on a continuous schedule of reinforcement (i.e., after each manding response); however, in naturalistic settings; manding is typically reinforced on an intermittent schedule of reinforcement. Lastly, although peer behavior was not measured in the study, the experimenters anecdotally noted that each peer began to emit the target mand during the participant’s 30-s access to the toy. As the peer did not receive the toy when he manded for it, this evoked some problem behavior associated with extinction, such as whining and crying.

Additionally, procedural fidelity was somewhat lower for Tommy (average = 80%) than it was for Jason and Noah (average = 95.2%). Although all three participants experienced increases in peer-directed vocal manding during intervention, it is possible that lower procedural fidelity during intervention for Tommy may have resulted in his lower per-
percentages of independent responding, as reflected in the middle panel of Figure 1.

Furthermore, as a measure of social validity, we asked participants and peers to point to a face to indicate whether they liked to share with friends and whether they thought sharing was being a good friend. While participants and peers provided favorable responses to these questions, given their ages, disabilities, and functioning levels, we cannot definitively determine whether their affirmative responses (i.e., circling a happy face) where in response to the social validity questions, or whether they were in response to some other aspect of the experimental, classroom, or personal context. For example, it is possible that participants’ and peers may have simply liked happy faces as opposed to neutral or frown faces, and thus circled them when presented with the option.

Finally, the intervention was implemented in both a separate room and the participants’ typical classroom in order to promote generalization of responses in the natural environment. However, as shown in Figure 1, we did not conduct baseline evaluations of participants’ levels of independent, peer directed manding in the natural classroom setting prior to intervention. Therefore, we cannot draw any definitive conclusions regarding whether the intervention actually produced generalization of learned responses across the two settings.

Future Research

The results of the study suggest several areas for future research. First, since challenging behavior and other social responses may have been affected by the intervention, but were not measured in the study, future research should include measurement of these important collateral responses. Second, the present study involved teaching children to mand for items physically within view. Since many naturalistic situations involve children manding for items that are not in view (i.e., pure mands), future research should seek to evaluate procedures for teaching children to engage in peer-directed mands for items that are not in view, or not in view for at least a portion of response opportunities. Relatedly, since mands were reinforced on a continuous schedule of reinforcement, researchers should seek to thin the schedule of reinforcement to an intermittent schedule that more closely approximates the density of reinforcement available within naturalistic settings, where manding seldom produces continuous reinforcement.

Young children with ASD may allocate their mands to adults because adult reinforcement tends to be more frequent and immediate. Rather that teaching children with ASD to mand with their peers, researchers could teach peers to be more effective listeners by providing immediate and frequent reinforcement for their peers’ mands, which may obviate the need for direct teaching procedures. Since peers in this study were always 1-2 feet apart, future research could evaluate teaching peer manding at varying distances, as would be the case in a typical classroom setting. Finally, as the study was conducted in a self-contained classroom exclusively with students with ASD, future research should seek to evaluate peer-directed manding with typical peers in inclusive settings, along with teaching broader communicative responses beyond the mand.

Conclusion

Peer-directed communication is a critical skill for young children with ASD. This study evaluated a teaching packing with differential reinforcement and least-to-most prompting to increase peer-directed, vocal manding with carrier phrases of young children with ASD within a single-case, multiple-probe design. Following intervention, three young children with ASD increased their peer-directed vocal manding with carrier phrases. Learned mands maintained up to four weeks after withdrawal of the intervention. Limitations of the study included lack of explicit teaching of, or measurement of, broader communicative responses beyond the mand. Future research should seek to evaluate procedures for teaching mands within more inclusive, naturalistic contexts, and to include measurement of, and teaching of, collateral social and communicative responses.
References


assessments: Is current practice state of the art?

Journal of Applied Behavior Analysis, 24, 189–204.


Received: 20 September 2018
Initial Acceptance: 14 November 2018
Final Acceptance: 15 December 2018
Comparison of Concrete and App-Based Manipulatives to Teach Subtraction Skills to Elementary Students with Autism

Laura Bassette  
Ball State University

Emily Bouck  
Michigan State University

Jordan Shurr  
Queen’s University

Jiyoon Park  
Michigan State University

McKenzie Cremeans  
Ball State University

Abstract: Elementary students with autism spectrum disorders (ASD) who struggle with mathematics may be limited in opportunities to pursue advanced mathematics potentially impacting post-school outcomes, including postsecondary education and employment opportunities. Interventions that provide visual instruction, such as manipulatives, along with systematic instruction may be beneficial in teaching students computational skills. The purpose of this study was to expand upon previous research comparing the use of concrete and app-based manipulatives to teach elementary students with ASD to solve subtraction problems. The results from this single case alternating treatment design study indicate students completed more steps independently per minute during the app-based manipulative condition. Accuracy improved during both intervention conditions, however, only two participants demonstrated improved maintenance scores. The implications of using manipulatives with students with ASD and areas of future research are discussed.

Students with autism spectrum disorders (ASD) may face challenges associated with learning mathematics, such as committing problem solving and processing mathematics language and symbols (Yakubova, Hughes, & Shinaberry, 2016). Teachers of students with ASD often identify mathematics as a challenging area (Titeca, Roeyers, Ceulemans, & Desoete, 2015). In comparison to typically developing peers, students with ASD tend to have below-average procedural calculation and challenges with basic number competencies (Bae, Chiang, & Hickson, 2015). Consideration of elementary level mathematical skills for students with ASD is critical given many of the procedural skills taught in elementary schools are important to access higher-level mathematics (e.g., problem solving; Spooner, Saunders, Root, & Brosh, 2017). Additionally, mathematical ability is critical to helping individuals with ASD be able to perform tasks relative to daily functioning, such as managing personal finances and counting money (Patton, Cronin, Bassett, & Koppel, 1997).

Previous research regarding mathematics instruction for students with ASD is limited, but expanding (Barnett & Cleary, 2015; Gevarter et al., 2016; King, Lemons, & Davidson, 2016). In their systematic review of the literature on mathematics instruction for students with ASD, Barnett and Cleary (2015) found six of the 11 articles published used visual representations, inclusive of manipulatives. In their review of mathematical interventions for students with autism, Gevarter et al. (2016) found five of the 13 articles used manipulatives to support teaching and learning with the majority involving treatment packages for students with ASD.
which focused on the mathematics and behavior. King, et al. (2016) also noted the prevalence of using packages to teach mathematics to students with ASD in their best evidence synthesis for mathematical interventions and it was reported studies focused more on functional mathematics or computational procedures.

**Manipulative-Based Interventions**

Concrete manipulatives are physical objects that can be held, manipulated, and used to teach specific concepts. Base Ten blocks are one type which illustrate numerical concepts in a tangible form (i.e., block units; Marley & Carbonneau, 2014). Base Ten blocks can be easily moved and re-grouped to teach place value and computational problem solving (Fuson & Briars, 1990). Teaching students to use concrete manipulatives was previously identified to support mathematical learning for students without disabilities (Marley & Carbonneau, 2014) and a meta-analysis identified a medium positive effect in 55 different studies (Carbonneau, Marley, & Selig, 2013). Furthermore, when students used concrete manipulatives to explore and master concepts, they were more engaged and motivated (Stein & Bovalino, 2001).

While concrete manipulatives are effective, the increased presence of technology in today’s classrooms supports an exploration of virtual manipulatives (Bouck, Chamberlain, & Park, 2017, Bouck, Shurr, Bassette, Park, & Whorley, 2018). These are similar to concrete, except they are images of blocks available through websites (i.e., internet-based) or tablets (e.g., app-based; Brasel & Gips, 2015). App-based manipulatives, conveniently available on devices (e.g., iPad), do not require internet. The images of the manipulatives are viewed within apps and are able to be rotated, flipped, and turned as if three-dimensional like concrete manipulatives. The interface of virtual manipulatives permits students to manipulate screen images similar to concrete blocks held in hand (Satsangi & Miller, 2017).

**Comparison of Concrete to Virtual Manipulatives**

The comparison of app-based and concrete manipulatives was previously explored for students with disabilities (e.g., Bouck et al., 2017, 2018). Bouck, Shurr, et al. (2018) compared the different manipulatives to teach middle school students with mild intellectual disability and learning disabilities to add fractions and found both types were effective in improving accuracy in solving problems with students displaying similar levels of independence in solving problems and task completion time. Bouck et al. (2017) also compared the different manipulatives to teach middle school students with mild intellectual disability and learning disabilities to solve subtraction problems and found both improved accuracy, however students were more independent and solved problems faster when using the app-based manipulatives.

For students with ASD, limited research also examined the effectiveness of concrete and app-based manipulatives. Bouck, Satsangi, Doughty, and Courtney (2014) compared the independence of elementary students with ASD when using concrete versus virtual manipulatives (i.e., Base Ten blocks available online via a computer) to solve subtraction with regrouping problems. Bouck et al. (2014) found the three elementary students with ASD were slightly more independent in solving the problems when using the virtual manipulatives; the students also reported a preference for the Internet-based virtual manipulative over concrete Base Ten blocks. Root et al. (2017) also reported the three elementary students with ASD and a moderate intellectual disability in their study examining systematic instruction and manipulatives performed better when using the virtual manipulatives, as compared to the concrete manipulatives and similarly reported participants with ASD preferred the virtual manipulative.

**Current Study**

Previous research exploring mathematical interventions for students with ASD suggests interventions that include: technological components (Bouck et al., 2014), behavioral strategies (e.g., prompting, reinforcement, modeling), and academic components (Gevarter et al., 2016) are effective. Preliminary research also suggests students with ASD may prefer app-based manipulatives and students with other disabilities solved subtraction prob-
lems more quickly when using these; however, efficiency in using manipulatives to solve subtraction problems was not explored in students with ASD. The purpose of the current study was to compare students’ accuracy and efficiency (i.e., steps completed independently per minute) when using the app and concrete manipulatives to solve subtraction problems. Additionally, this study explored the extent that manipulatives facilitated students’ ability to solve problems abstractly.

Method

Participants

Three elementary students with autism spectrum disorder (ASD), as identified by their school district, participated in the study. Inclusion criteria for participation included: (a) school diagnosis of ASD, (b) difficulties with grade-level subtraction skills as expressed by teacher or parent, (c) difficulties with grade-level subtraction skills as confirmed by the KeyMath™ numeration and addition and subtraction subtests, (d) ability to manipulate concrete blocks and app on an iPad, and (e) parental/guardian consent and student assent.

Arthur. Arthur was an 8-year-old, male, Caucasian, second-grade student. Arthur was eligible for special education services through his primary ASD diagnosis and secondary speech and language impairment. Arthur received his education primarily in the general education classroom with his general education teacher and was described in his individualized education program (IEP) as “a delightful young man to work with” who got along well with his peers. No IQ or standardized testing data was included in his IEP. In regards to math, it was noted that fluency was a challenge and previous curriculum based assessments (CBM) indicated he scored an average of 20/100 problems correctly on a 5-minute timed addition test. A math goal in his IEP stated, “After 36 weeks of instruction, Arthur will be able to fluently complete addition and subtraction facts. He will be able to solve at least 75 addition facts correctly.” Accommodations included: individual assessments, visual cues and prompts, rephrase or restate questions to allow for comprehension.

Reinforcement after sessions included watching YouTube videos.

Mark. Mark was a 10-year-old, male, Caucasian, third-grade student. He was eligible for special education services under the category of ASD and had a secondary language impairment. Mark received one hour of special education services in a resource room setting and was described by his general education teacher as “a fun kid with a great imagination.” A review of Mark’s IEP indicated he displayed difficulty in transferring skills to new activities and did best with a structured routine. No IQ, standardized testing data, or CBM data for math was included in his IEP. Math was reported to be a strength; however, Mark had an F in math at the time of the study and struggled with staying on-task. Mark participated in grade-level mathematics and accommodations during state assessments in math included: additional breaks, read test aloud, extended time, and tested in small group setting. A behavioral plan was also developed to address Mark’s motivation during academics. When working with the researchers, Mark enjoyed the one-on-one attention, but at times would have difficulty focusing. Mark reported an interest in Pokemon and therefore during training and subsequent sessions, the researchers engaged Mark in the activities by treating sessions as a game of Pokémon (e.g., Orange blocks were “Fire type” Pokémon). As reinforcement for participation, Mark watched Pokémon YouTube videos.

Jeff. Jeff was a 9-year-old, male, Caucasian, second-grade student. Jeff had a primary diagnosis of ASD and also had a secondary Language Impairment. Jeff received mathematics instruction in a pull out resource room setting. No IQ information was included in his IEP. Jeff was described in his IEP as having a “pleasant demeanor” and “cares about others.” Jeff tended to have difficulty in following multi-step and verbal instructions, as well as working independently. Jeff participated in the alternative state assessment and current NWEA scores indicated he was performing in the 10th percentile. Math objectives focused on selecting the correct operation when given a word problem and appropriate use of a calculator, manipulatives, or picture to solve computation problems. When working with the researchers, Jeff was typically compliant;
however, at times, he would struggle to get started on his math or stay on task.

Setting

All data sessions were conducted in quiet areas with minimal distractions. Specifically, participants worked one-on-one with researchers at a desk or table with 3–4 chairs available. Initial data (e.g., KeyMath assessment and baseline sessions) for Arthur and Mark were collected at their elementary school in the computer room. Since the school year ended before the study was completed, remaining sessions were conducted in quiet spaces available at the researchers’ university (e.g., the university’s Center for Autism Spectrum Disorders). Sessions for Jeff were conducted at the same university locations and also a quiet room located at his church.

The first author, who has expertise in implementing educational interventions for children with autism, trained all other researchers prior to data collection. Researchers were trained to mastery confirmed through in-vivo demonstrations. During sessions, researchers sat next to the participant to observe behavior and provide prompts. When interobserver agreement data (IOA) were collected, a second data collector was also present.

Materials

Materials for this study included researcher-constructed assessment sheets with five problems/sheet, pencils, task analysis data collection sheets with codes for the system of least prompts data, Base Ten concrete manipulative blocks, Base Ten Place Value Mat, and an iPad with the Base Ten Blocks app (Base Ten Blocks Manipulative by Brainingcamp [2017]). During sessions, participants were given a worksheet with the subtraction problems and a pencil.

The Base Ten concrete manipulatives used in the study included plastic blocks that represented ones (i.e., units), tens (i.e., rods), and hundreds (i.e., flats). The units were one squared cm, rods were 10 cm squared, and flats were 100 cm squared. A green Base Ten starter kit containing the various blocks along with an orange block set were used since these mirrored the colors in the Base Ten app. The blocks were three dimensional and provided a tangible representation of how many total units were in each block (e.g., the rod had the 10 total units). A place value mat printed on a sheet of paper was used only for concrete manipulative sessions and provided space for students to set up problems with spaces for the ones, tens, and hundreds. During concrete sessions, students placed blocks within the spaces on the place value mat to represent the numbers in the problems. The Brainingcamp Base Ten app was used for app manipulative condition. Within the app, there was a virtual place value mat (on the screen) where students could set up subtraction problems by using the virtual units, rods, and flats. The app is similar to the Base Ten blocks as it permits users to set up problems on the place mat but uses virtual representations (e.g., a virtual one unit block to represent 1). Subtraction problems were set up by adding blocks within the place value mat to represent the numbers in the problems. Base Ten blocks were manipulated through grouping, deleting, and clearing all from the tablet screen.

Experimental Design, Independent, and Dependent Variables

An adapted alternating treatment design was used to determine if a functional relationship existed between intervention and dependent variables (Sindelar, Rosenberg, & Wilson, 1985; Wolery, Gast, & Ledford, 2014). The Base Ten app and concrete blocks were the independent variables. Using the concrete blocks was defined as setting up the numbers for the subtraction problems with the blocks and then using the blocks to solve the problem. Using the app-based manipulative was defined as setting up the subtraction problem with the virtual blocks (i.e., adding in the blocks to represent the numbers in the problem the tiles) and then solving the problem. Participants were expected to record answers on the worksheets during all sessions.

The dependent variables for this study included (a) correct problems, defined as the percentage of subtraction problems answered correctly out of five; and (b) steps completed independently per minute, defined as the number of steps of the task analysis (see Table 1) completed independently per minute.
<table>
<thead>
<tr>
<th>Steps of Task Analysis for Subtraction Problem Type</th>
<th>Triple-Digit Minus Triple-Digit</th>
<th>Double-Digit Minus Double-Digit</th>
<th>Teens/Twenties Minus Single-Digit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CM: Set up the hundreds blocks for the first number of the problem. VM: Set up the hundreds blocks on the app for the first number of the problem.</td>
<td>1. CM: Set up the tens blocks for the first number of the problem. VM: Set up the tens blocks on the app for the first number of the problem.</td>
<td>1. CM: Set up the tens blocks for the first number of the problem. VM: Set up the tens blocks on the app for the first number of the problem.</td>
<td>1. CM: Set up the tens blocks for the first number of the problem. VM: Set up the tens blocks on the app for the first number of the problem.</td>
</tr>
<tr>
<td>2. CM: Set up the tens blocks for the first number of the problem. VM: Set up the tens blocks on the app for the first number of the problem.</td>
<td>2. CM: Set up the ones blocks for the first number of the problem. VM: Set up the ones blocks on the app for the first number of the problem.</td>
<td>2. CM: Set up the ones blocks for the first number of the problem. VM: Set up the ones blocks on the app for the first number of the problem.</td>
<td>2. CM: Set up the ones blocks for the first number of the problem. VM: Set up the ones blocks on the app for the first number of the problem.</td>
</tr>
<tr>
<td>3. CM: Set up the ones blocks for the first number of the problem. VM: Set up the ones blocks on the app for the first number of the problem.</td>
<td>3. CM: Set up the tens blocks for the second number of the problem. VM: Set up the tens blocks on the app for the second number of the problem.</td>
<td>3. CM: Set up the tens blocks for the second number of the problem. VM: Set up the tens blocks on the app for the second number of the problem.</td>
<td>3. CM: Set up the ones blocks for the second number of the problem. VM: Set up the ones blocks on the app for the second number of the problem.</td>
</tr>
<tr>
<td>4. CM: Set up the hundreds blocks for the second number of the problem. VM: Set up the hundreds blocks on the app for the second number of the problem.</td>
<td>4. CM: Set up the ones blocks for the second number of the problem. VM: Set up the ones blocks on the app for the second number of the problem.</td>
<td>4. CM: Set up the ones blocks for the second number of the problem. VM: Set up the ones blocks on the app for the second number of the problem.</td>
<td>4. CM/VM: Regroup ones or ungroup tens to solve the problem.</td>
</tr>
<tr>
<td>5. CM: Set up the tens blocks for the second number of the problem. VM: Set up the tens blocks on the app for the second number of the problem.</td>
<td>5. CM/VM: Count the number of tens and ones.</td>
<td>5. CM/VM: Count the number of tens and ones.</td>
<td>5. CM/VM: Count the number of tens and ones.</td>
</tr>
<tr>
<td>6. CM: Set up the ones blocks for the second number of the problem. VM: Set up the ones blocks on the app for the second number of the problem.</td>
<td>6. CM/VM: Count the number of tens and ones.</td>
<td>6. CM/VM: Write or state the answer.</td>
<td>6. CM/VM: Write or state the answer.</td>
</tr>
<tr>
<td>7. CM/VM: Ungroup/regroup 1 ten to solve the problem.</td>
<td>7. CM/VM: Write or state the answer.</td>
<td>7. CM/VM: Write or state the answer.</td>
<td>7. CM/VM: Write or state the answer.</td>
</tr>
<tr>
<td>8. CM: Subtract the ones (remove from board). VM: Subtract the ones (ghost).</td>
<td>8. CM/VM: Write the answer for the ones.</td>
<td>8. CM/VM: Write the answer for the ones.</td>
<td>8. CM/VM: Write the answer for the ones.</td>
</tr>
<tr>
<td>9. CM/VM: Write the answer for the tens.</td>
<td>10. CM/VM: Ungroup/regroup 1 hundred to solve the problem.</td>
<td>10. CM/VM: Ungroup/regroup 1 hundred to solve the problem.</td>
<td>10. CM/VM: Ungroup/regroup 1 hundred to solve the problem.</td>
</tr>
<tr>
<td>11. CM: Subtract the tens (remove from board). VM: Subtract the tens (ghost).</td>
<td>11. CM/VM: Write the answer for the tens.</td>
<td>11. CM/VM: Write the answer for the tens.</td>
<td>11. CM/VM: Write the answer for the tens.</td>
</tr>
</tbody>
</table>
Event recording using the permanent product of the written answers from worksheets were used to determine the former and event recording by way of direct observation was used to assess the latter. A stopwatch was used to record total duration.

**Procedure**

*Pre-assessment and baseline.* Student difficulty in solving subtraction problems was confirmed through performance on the KeyMath-3 pre-assessment. Following the KeyMath assessment, the types of subtraction problems students struggled with was confirmed with baseline probes. During baseline sessions, participants were given a worksheet and completed five subtraction problems via paper/pencil. A minimum of five baseline sessions was included for each participant and intervention began once a stable baseline was confirmed.

*Manipulative training.* Previous research indicates students need direct instruction to understand how to use manipulatives (Neil & Jarvin, 2011), therefore a training period was included following baseline. During training, participants learned how to use the Base Ten blocks and Base Ten app to solve problems. All participants completed concrete manipulative training first and then app training. Researchers provided explicit instruction first by modeling how to solve the problems using the manipulatives; then, through guided instruction where researchers prompted participants to solve the problems when they struggled or made an incorrect step (based on the associated task analysis sheet); and finally, by probing the participant’s ability to solve problems independently. Mastery criteria for training was based on independence data (i.e., task analysis of type of problem) during the problems solved independently phase of training. Initial mastery criteria was set at three problems completed 100% independently; however, this was modified partway through and ultimately, the criteria to progress from training to intervention required participants to solve three problems 80% independently or higher. For example, Arthur worked on double-digit subtraction problems which included a seven step task analysis. To move to intervention, Arthur needed to solve three different double-digit problems at the 80% independence criteria. This was met when he successfully completed a minimum of six (out of seven) steps independently on three different problems. If a participant did not reach mastery criteria, the training protocol was repeated.

During concrete training, participants learned to represent the different place values (depending on the type of problem they were working on) to set up the minuend (i.e., the top or first number) and subtrahend (i.e., the bottom or second number) in the problem. Participants were taught to ungroup from the furthest left column (i.e., hundreds or tens) and regroup in the associated column (i.e., tens or ones) to solve the problems by substituting a flat with 10 rods (for triple-digit problems only) or substituting a rod with 10 units (for triple and double-digit problems). Students were then taught to find the difference between the minuend and the subtrahend to perform the calculations. For example, when solving the problem 23-15, students would ungroup one tens block from 23 and add those 10 ones to the three ones, resulting in 13 ones. The student would then identify the numbers in the subtrahend (i.e., one tens block and five ones block). Students were then taught to align the subtrahend blocks up with the minuend blocks and then remove all identified blocks leaving only the remaining blocks (i.e., eight ones blocks in the minuend) for the answer. Lastly, students were taught to count the blocks that were left and record their answers on the worksheet.

Similar procedures were used to teach participants to use the app. Students were taught to set up the minuend (i.e., the top or first number) subtrahend (i.e., the bottom or second number) in the problem in the app. Students then learned how to drag a flat or rod from the left column over to the respective column (e.g., tens or ones) to ungroup and the app would automatically regroup the flat or rod when dragged to the new column. Students also learned to drag the blocks down in the app to perform the subtraction calculation. For example, if the problem was 22-16 the student would ungroup one tens block (from the 20) into the ones column creating a total of 12 ones blocks (the app would automatically ungroup the tens block into 10 ones
or the hundreds block into 10 tens blocks). The student would then drag down the tens block from the minuend (to take it away from the tens block in the 16) and the app would indicate they subtracted 1 out of 1 ten. The student would then move the 6 ones blocks (for the 6 from the 16) and the app would indicate when they subtracted 6 out of 6 ones. Then, only any blocks from the minuend remained on the screen which participants counted to get their answer. Once students wrote their answer, they learned to check their response by viewing the answer provided in the app.

**Intervention.** Intervention consisted of three alternating conditions: Base Ten concrete blocks, Base Ten app, and paper/pencil – no manipulatives (i.e., extended baseline). A total of 15 intervention sessions were conducted with five sessions of each of the three conditions. Session order was randomly assigned with no more than two consecutive sessions of the same condition. When the blocks or app was used, researchers implemented the system of least prompts if a student did not initiate the step within 10 seconds. Specifically, if a student did not perform a step independently, the following prompting scheme was used until correct performance began: gesture (e.g., “point to the number in the problem”), an indirect verbal prompt (e.g., “look at the tens value”), a direct verbal prompt (e.g., “count out three blocks in the ones place”), model (e.g., showing how to count out five blocks for the ones place), partial physical prompt (e.g., gently guiding the student’s hand to pick up an additional block to put on the place value mat), and full physical prompt (e.g., gently guiding the student to regroup a rod for 10 units).

**Concrete Base Ten blocks.** When students solved problems during the concrete sessions, they were provided with the worksheet with the five problems, a pencil, and the associated orange and green blocks based on the types of problems they were working on (e.g., students who were solving double digit problems were not given flats). For each problem, students needed to complete each step of the associated task analysis (see Table 1). For example, when solving double-digit minus double-digit problems, students needed to set up the tens for the minuend, set up the ones for the subtrahend, set up the ones for the subtrahend, regroup (i.e., tens columns), remove blocks, and write down the answer.

**App-based Base Ten blocks.** When students solved problems during the app sessions, they were provided with a pencil, the worksheet with the five problems, and the iPad with the Base Ten blocks app. Similar to the concrete sessions, students were expected to set up the numbers in the problems (e.g., hundreds, tens, and ones for the minuend and subtrahend for triple digit problems), regroup by dragging the associated rod or flat over, removing blocks to conduct the subtraction, and write down their answer on the worksheet.

**Paper/pencil (no manipulative).** During the extended baseline condition, participants were only provided with the subtraction worksheet problems and a pencil and no manipulatives (i.e., concrete blocks or app). Data collection was similar as to baseline.

**Best treatment.** The best treatment phase was each student’s best intervention condition – concrete or app – and procedures outlined by Wolery, Gast, & Ledford (2014) were followed. Specifically, consistent differences between data point values of the conditions alternated during the comparison phase were examined to determine if one condition was consistently superior when compared to the corresponding data point of the other conditions explored. The percentage of non-overlapping data (PND) for steps completed independently per minute data was used. PND was calculated by determining the number of sessions in one intervention condition (e.g., app) were superior to corresponding data points of the other condition (e.g., concrete) and dividing by the number of comparison sessions (five; Wolery et al., 2014). During best treatment sessions, students were given a pencil, a worksheet with five problems, and the associated manipulatives (i.e., Base Ten concrete blocks or Base Ten blocks app on the iPad). The system of least prompts was used in conjunction with the manipulatives during the best treatment sessions. Data of problems solved correctly and steps completed independently per minute were recorded using the same procedures used during intervention.

**Maintenance.** Baseline procedures were used to collect maintenance data 1–2 weeks
following completion of the best treatment phase. The students were only given the worksheet with five subtraction problems and asked to solve the problem and write their answers; no app, blocks, and prompting were not provided during this phase.

**Interobserver Agreement, Treatment Fidelity, and Social Validity**

Interobserver agreement (IOA) data for the percentage of problems solved correctly and independence was calculated for a minimum of 33% of sessions in each phase. IOA data for problems correct were calculated by a second researcher reviewing the permanent product of answers written on the worksheets and determining agreement of answers solved correctly or not. IOA data for independence, were recorded through direct observation of sessions during which a second trained observer simultaneously recorded any prompts recorded by the primary observer. The number of agreements of independence or the type of prompt delivered for each step of each problem was summed and divided by agreements plus disagreements. IOA data for duration data was recorded by the primary observer showing the second observer the time on the stopwatch when the session was completed and both observers recording the time. IOA was 100% for problems solved correctly for all participants. IOA independence for Arthur was 100% for best treatment and baseline sessions, and it was 98.6% for app and concrete sessions, 100% for Mark in all phases, and 100% for baseline and best treatment phases and 98.3% for app and concrete sessions for Jeff. IOA for duration data was 100% for all participants across all phases.

Researchers recorded treatment fidelity data for 40% of intervention sessions for each condition and 33.3% of best treatment sessions. Specifically, researchers monitored if students were given: the worksheet with the subtraction problems, the correct tool (i.e., concrete blocks, iPad with app, or no manipulatives), if students used the manipulative provided during the session. For each student for all phases and conditions, treatment fidelity was 100%.

Researchers conducted pre and post social validity interviews with the participants’ resource room teacher, each of the students, and the participants’ parents. The students were asked questions regarding which type of manipulative they liked the best, as well as if they preferred to solve the mathematics without any manipulatives.

**Data Analysis**

Visual analysis of the data was used to determine if a functional relation existed between the intervention and the dependent variables (i.e., problems solved correctly and steps completed independently per minute). A functional relation was assessed based on visual analyses of changes in level and trend, variability between phases and conditions, and fractionation of data between conditions. The researchers used Tau-U to calculate effect size for the accuracy data. Tau-U values represent contrasts between each intervention condition with baseline conditions for the same participant (Parker, Vannest, Davis, & Sauber, 2011). To calculate the Tau-U, researchers used a web-based calculator (see http://www.singlecaseresearch.org/calculators/tau-u; Vannest, Parker, & Gonen, 2011). Tau-U scores less than or equal to 65% indicate a small effect, 66–92% a medium effect, and 93% and above a large effect (Parker, Vannest, & Brown, 2009).

**Results**

Intervention effect was observed following training sessions where participants demonstrated increased ability to solve problems correctly (Figure 1). Fractionation of data for steps completed independently per minute indicated participants performed more steps on their own per minute during the app manipulative condition compared to the concrete condition (Figure 2). Participants also completed a higher number of steps independently per minute during the best treatment condition compared to intervention.

**Arthur**

On the KeyMath assessment, Arthur’s grade equivalency (GE) was 2.2 and Age Equivalency (AE) was 7:6; he struggled with accurately solving double-digit subtraction problems with regrouping (e.g., 42-29). During baseline, Ar-
thur did not solve any problems correctly and his data were stable, with a zero-eleration trend. Arthur required 19 independent training sessions to meet criteria for concrete training. Initial training criteria was three problems solved 100% independently (i.e., complete all task analysis steps without prompting). During concrete training, Arthur

Figure 1. Accuracy of subtraction problems.
completed 13 sessions with criteria set at 100% independence and he failed to meet the criteria in all sessions. Criteria for remaining concrete training (i.e., sessions 14–19), was modified to the 80% independent mastery criteria for three problems. App training fol-
Arthur

Arthur had used the three problems at 80% independence criteria for all sessions and he met criteria after seven sessions. During intervention, Arthur solved all of the problems correctly during the manipulative conditions, however his data were variable during extended baseline sessions (range 1–4, average 3). Fractionation of data was demonstrated in steps completed independently per minute; he completed more steps independently per minute during the app condition (range 2.3–3.5, average 2.5) than during the concrete condition (range 0.93–1.7, average 1.2). PND for app to concrete was 100% (concrete to app 0%) and the app condition was determined to be his best treatment condition. During best treatment, he completed all problems correctly and steps completed independently per minute ranged from 2.8–3.5, with an average of 3.2. Compared with baseline data, Arthur’s Tau-U for accuracy was 100% for both concrete and app-based manipulatives. Both of these indicate high effect sizes. The effects of the intervention were maintained as evidenced by an average of 4.67 of the problems solved correctly during maintenance.

Mark

Mark’s results on the KeyMath assessment indicated his GE was 2.5 and AE was 8:1 and he struggled with triple-digit minus triple-digit problems with regrouping (e.g., 800-692). During baseline, Mark solved an average of 0.4 of the problems correctly and a stable baseline was reached after five sessions with the last three sessions having a zero-celeration trend. Mark needed a total of 13 independent sessions to meet criteria for concrete training and initially, mastery criteria was set at completing three problems 100% independently. During concrete training, Mark completed 9 sessions with criteria at 100% independence and did not meet criteria during any sessions. Criteria for remaining concrete training (i.e., sessions 10–13), was modified to the 80% independent mastery criteria for three problems. App training followed concrete and the three problems at 80% independence criteria was used and he met criteria after three sessions. During intervention, Mark solved all of the problems during the app condition correctly, an average of 4.4 of the problems during the concrete sessions correctly, and an average of 2.6 of the problems during the extended baseline sessions correctly. Mark completed more steps independently per minute when using the app with an average of 4.8 compared to 3.2 steps completed independently per minute during the concrete manipulative condition. PND for app to concrete was 80% (concrete to app was 20%); the app was his best treatment condition. During best treatment, he solved all problems correctly and completed a higher number of steps independently per minute (range 6.0–8.4, average 7.1). Mark maintained an average of 4 problems solved correctly. Compared with baseline data, Mark’s Tau-U for accuracy was 80% for concrete and 90% for app-based manipulatives indicating medium effect sizes.

Jeff

Jeff’s results on the KeyMath assessment indicated his GE was less than kindergarten and AE was less than 4:6. Based on the assessment, he struggled with double-digit minus single-digit subtraction problems (e.g., 22-9). During baseline, Jeff solved 0 problems correctly and had a zero-celeration trend. Mastery criteria was set at three problems solved at least 80% independently during all of Jeff’s training sessions. Jeff met criteria for concrete manipulatives after nine sessions and after six sessions for app manipulatives. Once Jeff began intervention, he solved all of the problems correctly in both conditions. During the extended baseline phase, he solved an average of 0.2 problems correctly. Jeff completed more steps independently per minute during the app condition (range 1.5–1.9, average 1.7) than the concrete condition (range 0.6–1.7, average 1.2), suggesting fractionation of the data. PND for app to concrete was 100% (concrete to app was 0%) therefore, Jeff’s best treatment condition was the app condition. During best treatment, he solved all of the problems correctly and completed an average of 2.0 steps independently per minute. During the extended baseline and maintenance phases, Jeff solved one problem correctly. Compared with baseline data, Jeff’s Tau-U for accuracy was 100% for both concrete and app-
based manipulatives, indicative of high effect sizes.

Social Validity

Social validity questions were completed by participants and parents. Prior to intervention, participants indicated that math was an important skill and that they believed the iPad could assist them in learning new math skills. The parents and teachers indicated the app could be helpful as a teaching tool and that the students enjoy using the technology and they believed this could be useful in helping them learn new skills. After data collection, all three participants indicated they enjoyed the app more and felt it was easy to use to work on their math. Arthur stated he liked working on math using both the concrete manipulatives and the iPad but specified using the iPad “was awesome” and he would like to use it more at school. Mark said he liked that the app included graphics and it helped him to be able to figure out the correct answer. Jeff indicated he liked the app better because he preferred using the blocks on the screen.

Discussion

The purpose of this study was to expand the research exploring the use of app-based and concrete manipulatives to teach subtraction skills to students with ASD. The results suggest the treatments (i.e., manipulatives in conjunction with prompting) were effective in increasing students’ ability to solve problems accurately when using manipulatives and this may have assisted two of the participants’ ability to solve problems abstractly during maintenance. When comparing efficiency of manipulative type, the results indicate students completed more steps independently per minute when using the app-based manipulative. These results support previous research and provide additional evidence that both types of manipulatives are beneficial to supporting mathematics for students with ASD (e.g., Bouck et al., 2014, Root et al., 2017), and suggest app-based manipulatives may increase efficiency in solving subtraction problems (Bouck et al., 2017). Furthermore, this study expands upon previous research by including additional types of subtraction problems (i.e., teens/20s minus single digit and triple digit problems) which were not explored in the previous literature.

An evaluation of prompting required during intervention did not indicate a clear pattern of errors. For example, Arthur was prompted across various steps (e.g., regrouping, count the number of tens and ones, write the answer steps) during four of the app sessions and all concrete sessions (e.g., count the number of tens and ones during, set up the tens block for the first number during, regroup). Prompting for Mark also varied and was delivered during all five concrete sessions (e.g., subtract the ones, subtract the tens, ungroup/regroup one ten, set up the tens block, set up the hundreds block) and during four app sessions (e.g., ungroup/regroup one ten, write the answer for the ones, set up the hundreds block). During both concrete and app sessions, Jeff was prompted during the intervention sessions 1–3 and prompts varied across steps; however, no prompting was needed during the last two intervention sessions in either condition. During the app sessions that required prompting, he was prompted at least once on every step during at least one problem (e.g., set up the tens block, set up the ones block regroup) and during concrete sessions he was prompted at least once on every step except the final (i.e., write the answer).

While there is not a clear pattern of errors, the steps completed independently per minute data demonstrate fractionation between the app and concrete conditions with the app condition being superior for Arthur and Jeff (Mark to a lesser extent). Given the importance of procedural knowledge to access higher level mathematics, it is important to identify strategies to assist in teaching these skills. One way students demonstrate mathematical fluency is through displaying an understanding of symbol representation and being able to follow rules to complete mathematical processes (Van de Walle, Karp, & Bay-Williams, 2010). The increased efficiency in the app condition suggest components of the app-based manipulative may have assisted with procedural knowledge (e.g., the students’ ability to understand how the blocks represented the subtrahend and minuend and the
Implications for Practice

The results have several important implications for educators seeking to identify strategies to teach mathematical skills to elementary students with ASD. First, the results support previous research suggesting preference towards socially acceptable technologies (e.g., iPads) increases engagement of students with disabilities in academic activities (Bouck et al., 2014; Bouck et al., 2017). The students’ ability to use the app manipulative more efficiently may be a result of the interface. For example, regrouping in the app was displayed by dragging a rod or flat to the lower place value column (e.g., tens column) and once moved, the app automatically regrouped (i.e., broke up) the blocks. When using the concrete manipulatives, students had to manually count out new blocks to regroup. This required a greater response effort and, at times, resulted in errors if students did not count the blocks correctly. Also, at times during concrete sessions, participants displayed interfering behavior (e.g., fidgeted with the blocks) which appeared to function as way to escape from the task demands and likely contributed to the longer time needed to complete problems. The app interface also provided participants with additional visual information that was not available with the concrete manipulatives (e.g., the app would show the Arabic numerals in the app in addition to the blocks) and during training, participants were taught to attend to these cues. Educators using app manipulatives should ensure direct instruction is utilized to teach about all salient features of the app.

Students displayed greater efficiency in the app condition which may have assisted with their computational fluency (Carr & Alexeev, 2011). This is important given that a students’ ability to retrieve answers to computational problems fluently may serve as an indicator of their ability to understand abstract mathematical concepts (Jordan, Hanich, & Kaplan, 2003). Additionally, the app may be of assistance given that there is an option available for students to view the answer to the problem. This enables students to receive immediate feedback on their performance which could be helpful for teachers working with multiple students. Lastly, the Base Ten Blocks manipulative app is substantially cheaper compared to purchasing concrete manipulatives which may be beneficial for school districts or teachers who have iPads available.

Limitations and Future Directions

This study expanded upon previous work which compared app and concrete manipulatives to teach computational problems to students with ASD. A training period was included given previous research that indicates the importance of explicitly teaching students to understand manipulatives prior to use (Neil & Jarvin, 2011). Students were first trained to use the concrete manipulatives given that previous research suggested some students struggle with learning to use the blocks more than the app (e.g., Bouck et al., 2017) and anticipation that acquisition of the concrete manipulatives might take longer. One limitation during training was criteria varied for two of the participants and was decreased partway due to time constraints of the school. During initial concrete sessions (i.e., criteria of solving three problems 100% independently), researchers noted if students completed the problem entirely independently or not; however, prompting for each task analysis step was not recorded. Notably, Arthur and Marks’ best treatment condition was the app despite less exposure during training. Future research should explore acquisition of manipulative use under equivalent training. Additionally, it may be of interest to explore acquisition of concrete manipulative use following app training.

Another limitation of this study was the participants’ varying grade/ability levels. Mark was in third grade and completed grade level testing during the time of the study; however, Jeff and Arthur were in second grade. Arthur was on track to participate in grade level state standardized testing; however, Jeff was on track to participate in alternative testing suggesting academic were challenging for him. This could potentially be a contributing factor to his low performance during the maintenance phase where he did not complete any problems correctly. Future research that com-
pares the use of app-based manipulatives with students with ASD who have a more similar academic ability could assist in a greater understanding of how to best assist students with ASD of varying ability in maintaining ability to solve problems correctly once manipulatives are removed. Jeff also tended to rush through problems during the extended baseline and maintenance phase and appeared to lack motivation. Research that addresses motivation may help improve maintenance performance.

A third limitation is prompting was provided during intervention and best treatment phases; the impact of the manipulatives alone was not explored. Future research that incrementally reduces treatment components (e.g., prompts, visual cues) may provide insight into which components are critical to teaching students with ASD to ultimately transition to solving problems abstractly. This study provides preliminary evidence that some students with ASD may be more efficient in solving problems when using app-based manipulatives however. Future research exploring the component of time when solving problems with manipulatives is needed.

References


Patton, J. R., Cronin, M. E., Bassett, D. S., & Koppel,


Received: 25 July 2018
Initial Acceptance: 20 September 2018
Final Acceptance: 23 October 2018
Video Self-Modeling and Functional Behavior Assessment to Modify Aggressive Behaviors in Students with Autism Spectrum Disorder and Intellectual Disabilities

Kate M. Sadler
University of Virginia

Abstract: This paper presents a study that evaluated the efficacy of a video self-modeling intervention derived from a functional behavior assessment intervention on the aggressive behavior of two students diagnosed with autism spectrum disorder and intellectual disabilities. A multiple baseline design across stimulus conditions with inter-subject replication was conducted to explore the effects of the intervention. Data were collected on both the rate of aggressive behaviors and the use of replacement behaviors which fostered socially appropriate self-management. The data indicate an immediate decrease in aggressive behavior following intervention across all stimulus conditions for both participants. Data collected on the replacement behavior indicated a change from baseline to intervention but results differed across participants.

Video self-modeling (VSM) as a way to improve behavior has been in existence in the research literature for over 40 years (Creer & Miklich, 1970; Dowrick, 2012; Hitchcock, Dowrick, & Prater, 2003). In particular, much attention has been given to video modeling as a way to modify the behavior of students with autism spectrum disorder (ASD). Research has demonstrated that VSM meets the criteria as an evidence-based intervention for individuals with ASD, often producing substantial changes in student behavior without the need for intensive services (Bellini & Akullian, 2007; Bellini, Gardner, Hudock, & Kashima-Ellington, 2016; Dowrick, 2012).

Traditional video modeling consists of showing a person a video of someone performing a targeted skill whereas in VSM, the performer is the intended viewer (Bellini & Akullian, 2007; Cihak & Schrader, 2008; Gelbar, Anderson, McCarthy, & Buggey, 2012). While to some the model may seem a minor difference, VSM explicitly draws upon Albert Bandura’s research which suggests people learn best from models that most closely resemble themselves and having oneself as the model optimizes this factor (Bandura, 1977; Dowrick, 2012). Furthermore, the psychological implications of VSM suggests that viewers experience an increase in self-efficacy (i.e., the belief in one’s personal abilities to accomplish a goal) which correlates directly with success rates according to Bandura (1977, 1982, 2012).

VSM has been used as a school-based intervention for people with varying severity levels of ASD to modify many types of behavior such as: adaptive skills (Cihak, Fahrenkrog, Ayres, & Smith, 2010; Cihak & Schrader, 2008), compliance training (Buggey, 2005; Coyle & Cole, 2004; Lang et al., 2009), academic skills (Burton, Anderson, Prater, & Dyches, 2013; Hart & Whalon, 2012; Marcus & Wilder, 2009; Schatz, Peterson, & Bellini, 2016), social skills (Bellini & Akullian, 2007; Boudrea & Harvey, 2013; Buggey, Crawford, and Rogers, 2018; Buggey, Hoomes, Sherberger, & Williams, 2011; Buggey, Toombs, Gardener, & Cervetti, 1999; Hart & Whalon, 2012; Wert & Neisworth, 2003), and emotional recognition (Bernad-Ripoll, 2007). Studies suggest that the effects of a VSM are immediate and sustaining which are ideal for individuals with ASD who often
struggle with performance deficits and generalization (Bellini & Akullian, 2007; Buggey & Ogle, 2012, Gelbar et al., 2012).

VSM has clearly been successful, however there is very little research that uses VSM to address aggressive behaviors in students with ASD which reportedly occurs in over 50% of those diagnosed (Bronsard, Botbol, & Tordjman, 2010; Farmer et al., 2015; Mazurek, Kanne, & Wodka, 2013). Studies show among those with ASD, indicators for aggressive behaviors include intellectual disability (ID) and limited communication skills (Farmer et al., 2015; Goldin, Matson, Cervantes, 2014; Gotham, Brunwasser, & Lord, 2015). Among those diagnosed with ASD, approximately one-third also have an ID (Center for Disease Control, 2014). Often, for individuals with ASD and ID, maladaptive behaviors are highly entrenched, difficult to manage, occur at high rates, and tend to persist over the lifetime (Matson & Adams, 2014). Within the k-12 school setting, aggressive behavior is cause for the greatest loss of instructional time, reduced opportunities for independent functioning and interpersonal relationships, and restrictive placements for individuals with ASD (Farmer et al., 2015; Matson & Adams, 2014). These reports demonstrate that more work in this area is clearly needed (Matson & Adams, 2014).

Most school-based interventions to treat aggressive behaviors in individuals with ASD draw from principles of applied behavior analysis (ABA) which incorporates a wide range of strategies focused on observable interactions between the individual and the physical environment (Baer, Wolf, and Risley, 1968; Fisher, Piazza, & Roane, 2011). A functional behavior assessment is an ABA procedure used to gather information about a problem behavior and its likely reinforcing consequences through direct and indirect observations (e.g., interviews, questionnaires). In school settings, information from the functional behavior assessment is used to develop a behavior support plan that teaches a functionally equivalent replacement behavior often through a series of communication training or FCT (Carr & Durand, 1985) and/or reinforcement procedures (e.g., token economy). Literature provides strong evidence to support short term positive effects of function-based inter-
ventions (Campbell, 2003); however, for students with ASD and ID, there is a poorer prognosis on traditional behavior support plans (Kanne & Mazurek, 2011; Matson & Shoemaker, 2009; Rojahn, Wilkins, Matson, & Boisjoli, 2010).

For students who have made little to no progress on school-based behavior support plans a function-based VSM is ideal. Using a VSM to depict a functionally equivalent behavior offers an alternative to traditional procedures. For instance, a function-based VSM would provide the learner with a satisfying reinforcing consequence, visual portrayal of a communicative response amidst a condition that typically triggers aggressive behavior, and a point of view that has the ability to increase self-efficacy. Additional benefits would include the immediate and maintained effects that often accompany a VSM intervention.

At present, there are no studies that used a functional behavior assessment to develop target behaviors in a VSM and only one study that has focused specifically on reducing aggressive behavior. Buggey (2005) used VSM to reduce instances of pushing displayed by a pre-school student who was moderately affected by ASD. The video depicted the student demonstrating appropriate behavior amidst a condition that typically triggers aggression. Buggey (2005) reported that after viewing the video one time, the child’s pushing behavior occurred only once and no incidence was reported during the rest of the intervention or maintenance condition. The results of this study are astounding considering the minimal invasiveness (e.g., delivered one time daily) and immediate behavioral changes. Similar to other school-based interventions for aggressive behavior, the VSM intervention redirected an aggressive behavior with a socially appropriate behavior (e.g., standing in line with hands to self).

Despite promising results, limitations to this study need to be acknowledged. First, Buggey (2005) treated aggressive behaviors without the use of a functional behavior assessment. Therefore, the behavior function was merely assumed. Without a hypothesis of behavior function, it is not possible to ascertain that the target behavior depicted in the video eliminated the aggression. For instance, it is possible that the novelty of the intervention de-
increased the behavior in the depicted situation, but the function or satisfying consequence of the aggression was still unmet. If the target behavior in a VSM is derived from a functional behavior assessment, the likelihood that it is consequentially satisfying for the learner is high. Second, as there was only one participant in one condition (i.e., lining up for recess) further replication is required to determine if this is a viable means to modify aggressive behavior in other settings/conditions.

**Purpose of the Current Study**

The purpose of this study was to test an intervention that uses a functional behavior assessment to determine the target behaviors (e.g., replacement behaviors) featured in a VSM. It is possible that the strengths of these techniques can be combined to create an effective intervention that could possibly overcome high rates of persistent aggression, poor prognosis in behavior support plans, and add to the literature base for individuals with ASD and ID. The following central question and two sub-questions guide this study: To what degree does a VSM intervention developed from a functional behavior assessment modify classroom behavior in students with ASD and ID? Sub Aim 1) What effect will the VSM have on aggressive behavior?; Sub Aim 2) What effect will the VSM have on the replacement behavior?

**Method**

**Setting**

Two students, Nick and David (pseudonyms) participated in the study. Both students attended a self-contained public school in the Midwestern United States. The school served approximately 172 individuals in a sub-urban area, ages 5 to 14 years, with a variety of moderate and significant special needs as identified by educational diagnosis. The primary form of instruction in the school utilized Applied Behavior Analysis (ABA) principles (e.g., behavior intervention plans, functional assessments, structured environmental supports, and functional communication systems) in addition to school-wide positive behavior supports (SWPBIS), and state curriculum standards.

**Participants**

The first participant, David was 7 years old and in first grade at the time of the study. According to school records David had a score of 42 on the Childhood Autism Rating Scale™-2 (CARS-2; Schopler, Van Bourgondien, Wellman, & Love, 2010), which classified him as having a “severe” ASD. He had a score of 42 on the Developmental Profile™-2 (DP-2; Alpern & Boll, 2007), which classifies him as having an intellectual disability. David had been on his current behavior plan for 9 months and it listed “access to tangible items” as the main function of behavior and he was taught to make verbal or gestural requests (e.g., pointing) and participate in a self-management technique (e.g., counting to 10, deep breaths). Despite educational supports and high levels of reinforcement (e.g., verbal praise, access to preferred items) his teacher reported high levels of aggressive behavior which included: throwing objects, scratching, hitting, biting, and pulling hair of those around him.

Participant two, Nate was 12 years old at the beginning of the study. According to school records he had a score of 76 on the Social Responsiveness Scale™ (SRS; Constantino & Gruber, 2012) which classified him as having “severe” ASD. He had a score of <40 on the DP-2 which classified him as having an intellectual disability. Nate had been on his current behavior plan for 11.5 months, it listed “escape” as the main function of behavior and he was taught to request “a break” and participate in a self-management technique (e.g., squeeze a ball, pull therapy bands). Despite educational supports and high levels of reinforcement (e.g., verbal praise, access to preferred locations) his teacher reported high levels of aggressive behavior which included: biting, kicking, pulling hair, and scratching.

**Pre-Baseline Measures**

Several measures were implemented to determine student preferences and the function of aggressive behavior in order to develop the VSM intervention.
**Preference assessment.** Each child’s caregiver and classroom teacher completed an interest inventory and a questionnaire to determine communication style and preferences for the participants. From the data, it was determined that David communicated verbally using 2–3 word phrases to communicate his wants and needs and was academically working at a pre-kindergarten level. David enjoyed navigating YouTube on the computer, watching characters from Dinosaur Train, and playing with toys that he brought from home. Nate communicated his wants and needs using 1–2 words on a Talk Tablet® communication device. Academically, he was working at a kindergarten level. Nate enjoyed Dora the Explorer, listening to music, and lying down on a bean bag chair.

**Aberrant behavior checklist.** Students were evaluated using the Aberrant Behavior Checklist (ABC-2; Aman & Singh, 2017; \( \alpha = .65 \) to \( .41 \)) to determine the severity of their aggressive behaviors and offer a hypothesis of behavior function. Scores indicated that David exhibited high rates of maladaptive behaviors as an attempt to seek attention and escape from task demands during instruction. Scores indicated that Nate exhibited high rates of mal-adaptive behaviors as an escape from demands.

**Functional behavior assessment.** A functional assessment of behavior was conducted to ensure that the reinforcing consequence of the aggressive behavior was precisely targeted. Participants underwent a functional assessment of behavior by a Board Certified Behavior Analyst (BCBA). For this study two data sources, Questions about Behavioral Function (QABF; Matson & Vollmer, 1995; \( \alpha = .88 \) to \( .79 \)), and Antecedent-Behavior-Consequence (ABC) were utilized by the BCBA certified observer.

**QABF.** The BCBA observer gathered data from two teachers per participant (i.e., David’s teacher and para-professional and Nate’s teacher and school behavior specialist). Results from the QABF indicated that David’s behaviors were equally split between a desire for attention and an effort to escape a task demand. Results from the QABF indicate that Nate’s aggressive behaviors were the result of a desire to escape a task demand.

**ABC observations.** Based upon the results from the QABF, specific antecedents were targeted, and anecdotal consequence data was recorded for each participant. According to the data, the most commonly occurring antecedent to David’s aggressive behavior was the presentation of demands. Gaining adult attention was also identified as a possible function in a limited number of instances. For Nate, the most commonly occurring antecedent to aggressive behavior was the presentation of demands.

**Hypothesis statement.** Based on the data collected from the functional behavior assessment it was hypothesized that David’s aggressive behavior served two functions (escape a task demand and to seek attention) based on different environmental stimuli. Nate’s aggressive behavior served to escape task demands. Due to these results, it was determined that both David and Nate would be taught to request a break from demands and to participate in an individualized self-management strategy.

**Research Design**

For this study a multiple baseline design across stimulus conditions with inter-subject replication was used (Ledford & Gast, 2014; Horner et al., 2005). A stimulus condition was defined as an instructional arrangement (e.g., one-on-one instruction, group instruction, or a particular classroom setting). Three unique stimulus conditions were identified per participant based upon data obtained from preference and functional assessment measurements.

**Independent Variable**

The independent variable was the VSM. The fundamental aspects of the independent variable was consistent with previous VSM studies: (1) self-as-model; and (2) features the individual independently performing the target behavior under typically stressful conditions (e.g., when given a non-preferred task demand). Information from the preference assessments was used to ensure that the videos were at the appropriate communication and developmental level for each participant. David’s video was placed it on a desktop com-
puter and Nate’s was put onto a personal iPad. Classroom staff was trained on how and when to play the videos. A checklist was completed by the primary investigator to ensure the intervention was being delivered with fidelity.

**Video development.** Video footage was captured (using an iPhone 6 plus) and edited by the primary investigator to create a systematic, engaging movie using Windows Movie Maker® and Screencast-O-matic® software. Sounds effects (e.g., hand clapping) and positive praise (e.g., “great job!”) were incorporated into both videos. A background song, unique to each participant, was played on low volume for the duration of the video. For David, the VSM featured him being presented with a non-desirable task, requesting a break, participating in a self-management strategy (e.g., 10 deep breaths), being represented with the original task and upon completion, receiving a desirable consequence amidst three problematic stimulus conditions: (1) an academic task at his desk, (2) in the lunch room, and (3) in the Music classroom. Nate’s video featured him being presented with a non-desirable task, requesting a break, participating in a self-management strategy (e.g., squeezing a ball), being represented with the original task, and upon completion, receiving a desirable consequence amidst three problematic stimulus conditions: (1) small group reading, (2) science/social studies instruction, and (3) individual work at his desk.

**Dependent Variables**

**Aggression.** The first dependent variable was aggressive behavior and was the same across participants. For the purpose of this study, “aggressive behavior” was operationally defined as any instances of the following: physical contact that results in harming oneself or another individual (e.g., biting, hitting, kicking, scratching, pulling hair, or using items as weapons).

**Replacement behavior/self-management technique.** The replacement behavior was operationally defined as any time the student requested a break and participated in a self-management technique. To request a “break” from demands, David stated the word “break” and then proceeded to count to 10 and take 10 deep breaths. Nate used a picture icon or activated his Talk Tablet® device to request a break. Afterwards, he selected either a ball to squeeze or a resistance band to pull.

**Data Collection**

A partial interval recording system was used to collect data on the two dependent variables (aggression and replacement behavior) for a total of 25 non-sequential days. The partial intervals were 10 seconds in length. An observational period lasting 30 minutes occurred for each stimulus condition per participant. This allowed for a total of 180 intervals per stimulus condition for a total of 540 intervals of data collection per participant each day. Data were collected by the primary researcher during every interval, and an independent observer simultaneously recorded student behavior for 40% of the total number of intervals.

**Baseline.** Baseline conditions consisted of typical classroom interventions. The baseline condition continued until a pattern in the data for the dependent variables was apparent (Horner et al., 2005). Once a baseline pattern in the data was established, the intervention was presented in the first condition, while baseline procedures continued in the second and third conditions. When a pattern of data was established in the first condition, the intervention was then presented in the second condition, while baseline procedures continued in the third condition. When a pattern in the second condition occurred, the third condition began the intervention phase (Ledford & Gast, 2014).

**Stimulus condition 1 (SC1).** Following baseline, the VSM depicting the student engaging in the replacement behavior during SC1 was shown one time at the beginning of the school day. For David, the VSM for this condition depicted him using his replacement behavior after being given the prompt to engage in a reading task in his classroom. Nate’s VSM for this condition depicted him using his replacement behavior while engaging in small group reading instruction in his classroom.

**Stimulus condition 2 (SC2).** Once a pattern in the data was established in SC1 (e.g., three to five data points) the VSM that depicted the student engaging in the replacement behavior during both SC1 and SC2 was then shown at the beginning of the school day. For David, SC2 added a depiction of him using his re-
replacement behavior while eating in the lunchroom. For Nate, SC2 added footage of him using his replacement behavior after being given a prompt to engage in science/social studies curriculum (located in a semi-private desk in the back of the classroom).

Stimulus condition 3 (SC3). Once a pattern was established in SC2, the VSM that depicted the student engaging in the replacement behavior during SC1, SC2, and SC3 was shown at the beginning of the school day. For David, SC3 added a depiction of him using his replacement behavior during activities in the Music classroom. For Nate, SC3 added footage of him using his replacement behavior after being given a prompt to engage in a “task box” (e.g., sorting task) while he was seated at his own desk inside the classroom.

Maintenance. For David, maintenance data was collected on days 5, 9, 17, 20, and 24 post-intervention. For Nate, maintenance data was collected on days 3, 6, 10, 18, and 21 post-intervention. During the maintenance phase, the teacher and behavior analyst were instructed to cease playing the videos for the student each morning, but were asked to honor any request independently made by the student to view the VSM (e.g., if the student sought adult assistance to watch the VSM).

Inter-Observer Training and Agreement
Prior to collecting data, independent observers participated in a training session using videos that consisted of samples of both aggressive behavior and replacement behavior per participant. During training, inter-rater scores were calculated and the training(s) continued until 100% agreement was achieved (e.g., the data collectors came to an agreement that an aggressive or socially appropriate behavior did or did not occur). To practice interval synchronicity, observers were trained to listen to a private audio recording (using Smartphone technology) that indicated the beginning and end of each interval intended for observation.

Data were collected by the primary researcher 100% of intervals. IOA data were collected on 40% of the total number of intervals by a trained independent observer (e.g., doctoral students in the college of Special Education). This percentage is sufficient according to standards in the literature (Horner et al., 2005; Kratochwill et al., 2012). The IOA was evenly dispersed among baseline, intervention, and maintenance phases. The total number of agreements divided by the total number of agreements and disagreements yielded the percentage of inter-observer agreement. Inter-observer agreement for all collected data was at 99%. A Kappa statistic was established by dividing the total agreement by the chance agreement. Cohen’s Kappa was calculated at 0.98 which indicates almost perfect agreement.

Social Validity
Social validity was measured via a researcher-developed open ended questions given to two classroom teachers and two paraprofessionals familiar with the participants. The interviews allowed the primary investigator to ensure that the VSM intervention is minimally invasive to classroom instruction and helped determine the feasibility of dissemination. The social validity questionnaire consisted of the following questions: (1) Are you able to deliver the intervention with fidelity?, (2) Do you have the appropriate resources to deliver the intervention?, (3) Do you feel that the self-regulation strategy has social importance in the school?, (4) Do you believe the intervention is having an impact on the student?, (5) Would you consider using the intervention after the study is complete?, (6) Do you are any of your staff require additional instruction or support regarding the intervention procedures?

Results
A visual analysis of the data was performed to determine whether evidence of a functional relationship between an independent variable and an outcome variable exists and the strength or magnitude of that relation (Kratochwill et al., 2012; Ledford & Gast, 2014). Figures 1 and 2 present the instances of the dependent variables across stimulus conditions. Instances of aggression per interval are indicated in black with closed circles. Instances of the replacement behaviors per interval are indication in gray with open squares. The x-axis represents work sessions and the y-axis represents the frequency of the
dependent variables. Tables 1 and 2 represent mean performance evaluations and effect size calculations for both participants.

David

Baseline. David displayed variable rates of aggression (range 1 to 45 instances) and replacement behaviors (range 0 to 18 instances) during baseline conditions. During SC1 aggressive incidents occurred on average 33.5% of intervals; SC2 at 13.88%; SC3, at 11%. During SC1 the replacement behavior occurred on average 9.75% of intervals; SC2 at 4.5%; SC3 at 4.46%.

Intervention. During intervention, all three stimulus conditions demonstrate decreased levels of aggression immediately after the intervention is introduced. Interestingly, the level of replacement behavior is either higher than or equal to the rate of aggression across all stimulus conditions, with the exception of data point 11 in SC2. Both dependent variables demonstrate stable rates of behavior compared to the variable baseline conditions, suggesting that the intervention brought the behaviors under some degree of control. During SC1 aggressive incidents occurred on average 1.31% of intervals; SC2 at 0%; SC3, at 1%. During SC1 the replacement behavior occurred on average 3.75% of intervals; SC2 at 1.9%; SC3 at 2.16%.

Maintenance. During the maintenance condition, rates of aggression and replacement behavior mimic intervention conditions. Aggressive incidents occurred on average 0% of intervals during all three stimulus conditions. During SC1 the replacement behavior occurred on average 3.6% of intervals; SC2 at 3.4%; SC3 at 3.2%.

Nate

Baseline. Nate displayed variable rates of aggression (range 1 to 55 instances) and consistent low rates of replacement behaviors during...
baseline conditions. When presented with academic tasks during SC1 aggressive incidents occurred on average 25.8% of intervals; SC2 at 19%; SC3, at 17.83%. During SC1 the replacement behavior occurred on average 0% of intervals; SC2 at 4%; SC3 at 2.23%.

**Intervention.** During intervention, all three stimulus conditions demonstrate a pattern of

**TABLE 1**
David and Nate’s Mean Occurrence of Dependent Variables across Stimulus Conditions

<table>
<thead>
<tr>
<th></th>
<th>Aggression</th>
<th></th>
<th>Replacement Behavior</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Intervention</td>
<td>Maintenance</td>
<td>Baseline</td>
</tr>
<tr>
<td>David</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC1</td>
<td>33.50%</td>
<td>1.31%</td>
<td>0.4%</td>
<td>9.75%</td>
</tr>
<tr>
<td>SC2</td>
<td>13.88%</td>
<td>0.72%</td>
<td>0.2%</td>
<td>4.50%</td>
</tr>
<tr>
<td>SC3</td>
<td>11.00%</td>
<td>1.00%</td>
<td>0.4%</td>
<td>4.46%</td>
</tr>
<tr>
<td>Nate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC1</td>
<td>25.8%</td>
<td>0.53%</td>
<td>0.2%</td>
<td>0.2%</td>
</tr>
<tr>
<td>SC2</td>
<td>19%</td>
<td>0.20%</td>
<td>0.8%</td>
<td>4.0%</td>
</tr>
<tr>
<td>SC3</td>
<td>17.83%</td>
<td>0.80%</td>
<td>0.0%</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

*Note.* SC1 = Stimulus condition 1, SC2 = Stimulus condition 2, SC3 = Stimulus condition 3.
decreased levels of aggression and increased levels of replacement behaviors immediately after the intervention is introduced. Patterns of aggression are stable compared to the variable rates during baseline conditions. Replacement behaviors occur at an elevated rate compared to baseline conditions. Aggressive incidents occurred on average 0% of intervals during all three stimulus conditions, strongly suggesting that the intervention had the desired effect on aggressive behaviors. When presented with academic tasks during SC1 the replacement behavior occurred on average 12.8% of intervals; SC2 at 10.06%; SC3 at 33.2%.

**Maintenance.** During the maintenance condition, rates of aggression and replacement behavior mimic intervention conditions. Aggressive incidents occurred on average 0% of intervals during all three stimulus conditions. During SC1 the replacement behavior occurred on average 33.2% of intervals; SC2 at 9.4%; SC3 at 12%.

**Effect Size**

The percentage of data points exceeding the median (PEM) (Ma, 2006) was used to determine the effect size of the intervention. PEM was selected because the aggressive behavior was highly variant for these participants and a single ceiling or floor level data point would not be indicative of actual behavioral performance (Ma, 2006; Vannest & Ninci, 2015). In this type of analysis, the median of data points during the baseline phase are calculated and compared to the intervention and maintenance phases to determine the effect size of the intervention (Ma, 2006). The following guidelines were used to explain the effects using PEM: 90–100% is a highly effective treatment, 70–89% is a moderately effective treatment, 50–69% is a minimally effective treatment, and less than 50% is not effective (Scruggs, Mastropieri, Cook, & Escobar, 1986).

Table 2 displays the PEM calculations of the dependent variables (aggressive and replacement behavior) across stimulus conditions for David and Nate. For aggression, PEM rates across all stimulus conditions and phases (intervention and maintenance) were below the median rate of baseline aggression, suggesting that the intervention was highly effective both participants. For replacement behaviors, the data suggest no intervention effect David, whereas differing effects occurred for Nate. The PEM rate between baseline and both intervention and maintenance conditions in SC1 suggest a high intervention effect. The PEM rate between baseline and intervention in SC2 suggest a mild effect, whereas the PEM rate between baseline and maintenance in SC2 suggests no effect. For SC3, the PEM rate between baseline and both intervention and maintenance conditions suggest a high intervention effect.

**Social Validity**

For the both participants, the special education teacher and the student’s paraprofes-
sional were interviewed to determine if the intervention was socially valid. In regard to social importance, Nate’s teacher stated that, “he needs to learn to participate in groups and we finally found something that works.” For David, the paraprofessional stated, “Other staff was like whoa, what he is doing in the cafeteria?” She continued by saying, “They kept complimenting me on his behavior and I was like, it’s not me it’s the video!” When asked about the effect the video had on the Nate, his teacher reported, “I noticed a huge decrease in his behaviors and his willingness to participate in activities.” David’s special education teacher reported that she had noted a huge difference in his behaviors and that it is “exciting” to have David participate in the curriculum. All responders agreed that they would like to continue some form of VSM after the study was complete, but admitted that the technology seemed hard to use.

Discussion

The current study used multiple baseline design across stimulus conditions with inter-subject replication to explore the following central question: To what degree does a VSM intervention developed from a functional behavior assessment affect classroom behavior in students with ASD and ID? Two sub-questions also guided this research: (1) What effect will the VSM have on selected aggressive behavior?; (2) What effect will the VSM have on selected replacement behavior? To answer these questions, two students with ASD and ID who displayed high levels of aggressive behavior across various stimulus were identified to participate in the study.

VSM Effect on Aggressive Behavior

The results of the study provide strong support for the effectiveness for the VSM intervention derived from a functional behavior assessment for reducing aggressive behavior of students with ASD and ID. Both visual analysis and effect data demonstrated an immediate decrease in aggressive behavior across all stimulus conditions for both students. For David, aggressive behaviors ranged from 11–33.5% during baseline and ranged from 0–1% during intervention. For Nate, aggressive behaviors ranged from 17.83–25.8% during baseline and remained at 0% during intervention. Visual analysis indicated that the behavior change occurred rapidly after the onset of intervention and that this pattern was present for the duration of the intervention phase. Both participants maintained low levels of aggression (0%) after the removal of the intervention. These results add to the limited research of identifying an evidence-based intervention specifically designed to treat aggressive behaviors for students with ASD and ID (Matson & Adams, 2014; Matson & Shoemaker, 2009).

Using a functional behavior assessment to develop target behaviors in a VSM could be particularly advantageous for students with persistent aggressive behavior who do not respond to typical behavior modification strategies (Gelbar et al., 2012). For instance, both participants were taught a replacement behavior and self-management strategy in accordance to ABA methodology however, these strategies where demonstrated inconsistently and high rates of aggressive behavior continued. When the VSM was introduced, the replacement behavior and self-management strategy brought the behavior under some degree of control as indicated by dramatic decreases in aggression.

Equally important, the findings are consistent with previous VSM research which found that the intervention effects often occur immediately and maintain over time (Bellini & Akullian, 2007; Dowrick 2012; Gelbar et al., 2012). For both participants, decreases in aggressive behavior were immediate following intervention across all stimulus conditions. Furthermore, the effects of the intervention lasted up to 24 days for David and up to 21 days for Nate. These types of outcomes are exactly what one would hope for in a school-based intervention for aggressive behavior.

VSM Effect on Replacement Behavior

The results of the current study provide moderate support for the intervention’s potential to increase socially appropriate replacement behaviors of students with ASD and ID. For Nate, there was a strong effect size related to the VSM for replacement behavior, however for David, the effect size calculations suggest
that the VSM had no effect on the frequency of the placement behavior. David’s visual analysis suggest, however that there may have been some effect due to changes in the response patterns between baseline and interventions conditions.

A possible explanation for David’s mixed results could be his response to naturally maintaining contingencies in the physical environment. Often during behavior modification once aggressive behavior subsides, students are introduced to naturally occurring reinforcing consequences that had previously been overlooked (Baer et al., 1968). For David, simply viewing the reinforcing consequences in the video (e.g., praise and access to preferred items) may have helped him appropriately regulate his behavior, rendering the replacement behavior unnecessary. It’s equally possible that Nate internalized that replacement behavior, but required the use of items that provide sensory input to regulate his behavior. Further evidence of this internalization effect is apparent during the maintenance condition when the videos ceased from being played but patterns consistent with the intervention condition continued for both participants. The results are encouraging for both participants and warrant further investigation.

Limitations of the Current Study and Suggestions for Future Research

Several limitations may have affected the overall interpretations of this study and future research is needed. First, similar to other single-subject designs a small sample size was examined. Further, this study was the first of its kind; therefore, the results are in need of replication and should be interpreted with caution. Additional studies should consider larger samples and multiple sites as a way to determine the generalizability of the results. Next, maintenance data was only collected up to one month post intervention. Whether or not the effects of the intervention were maintained beyond that date is not possible to determine. Future studies should examine the effects over a longer duration of time. Additionally, in the current study the same function (e.g., task avoidance) was targeted for both participants. The effects of this intervention should be tested on other functions for aggressive behaviors such as attention seeking behaviors, behaviors seeking accessing to materials, and sensory seeking behaviors.

Moreover, it is unclear why the VSM has no effect on David’s replacement behavior. As noted, it is possible that the decrease in aggressive behavior could indicate that David learned to appropriately regulate his behavior. Due to lingering questions, additional studies should be carried out to analyze rates of replacement and aggressive behaviors. Future research should find ways to record responses to naturally occurring reinforcing stimuli and measure an individual’s self-regulation patterns pre/post an intervention to treat aggressive behavior.

Implications for Practice

There are several potential implications for practice that emerged from this study. First, teachers need interventions that are not time consuming and immediate in their results. Literature supports that in virtually all VSM studies the intervention produced results that accelerated quickly from baseline performance and were maintained in follow up assessments (Bellini & Akullian, 2007; Dowrick, 2012; Gelbar et al., 2012). Of particular importance is that the VSM maintenance effects can withstand frequent interruptions common in schools (e.g., holiday breaks, schedule changes or interruptions, and/or absence from school) (Bellini & Akullian, 2007; Hitchcock et al., 2005). Second, functional behavior assessments are widely used to identify the cause of aggressive behaviors in classroom settings. Teachers could use the existing information (e.g., data, hypothesis of behavior function) to determine which behavior should be targeted in a VSM intervention. Finally, given that many students are using technology in the classroom the use of the VSM intervention may be a socially valid approach to use in the classroom setting.

Conclusion

Individuals with ASD and ID have a poorer prognosis amongst traditional behavior intervention plans compared to those with ASD-only (Goldin et al., 2014; Matson & Adams,
Aggression is problematic for caregivers and teachers and can become debilitating for the individual. Based on the results from this study, VSM and functional behavior assessment has the potential to modify both aggressive behavior and replacement behavior within a classroom setting by addressing external reinforcing consequences and internal mechanisms that could account for sustained maladaptive behaviors.

References


Received: 15 August 2018
Initial Acceptance: 11 October 2018
Final Acceptance: 20 November 2018
Inclusive Social Studies Content Instruction for Students with Significant Intellectual Disability Using Structured Inquiry-Based Instruction

Joanna Ryan, J. Matt Jameson, Olivia Fudge Coleman, Carrie Eichelberger, Jessica A. Bowman, Lyndsey Aiono Conradi, Susan S. Johnston, and John McDonnell
University of Utah

Abstract: A high-quality education in the social studies is recommended by the National Council for the Social Studies to promote civic competence and democratic participation for school system graduates. Currently, limited empirical information exists to inform social studies instruction for students with significant intellectual disability (SID), especially within inclusive academic settings. Effective content instruction in social studies areas, aligned to the general academic curriculum and presented in the general setting, may help to improve post-school outcomes for students with SID. A multiple probe across participants single case design was used to determine the effectiveness of paraprofessional-implemented inquiry-based social studies instruction, presented within an embedded trial distribution schedule in general education classrooms. Study results suggest that this method of instruction had a positive effect on recall of information presented during history lessons for three junior high school students with intellectual disability. After repeated practice, students showed improvements in recall of information when using the structured inquiry-based instructional framework without direct instruction. Paraprofessionals and general education teachers generally found the intervention to be effective and acceptable. Implications for classroom practice and for future research are discussed.

In a recent position statement regarding teaching and learning in public schools, the National Council for the Social Studies (NCSS; 2016) states that civic competence depends on an excellent education in the social studies. If a purpose of public schooling is to promote civic competence for students, then this statement implies that schools bear a responsibility to provide high-quality instruction in social studies content areas so that graduates may successfully participate in society during adulthood. This may be an especially important consideration for educators who support students with significant intellectual disability (SID). Students with SID experience significantly diminished post-school outcomes in the areas of education, employment, and community access and participation (Bouck, 2012). Historically, educational planning for students with SID has focused on the instruction of functional skills in educational settings separated from the general student body (Hunt, McDonnell, & Crockett, 2012). Some special education researchers are currently examining ways to deliver instruction for students with SID that focus on access to, and progress in, the grade-level general education content within inclusive general educational settings (Ryndak, Jackson, & White, 2013). This mode of providing educational service, if leveraged in a way that provides individual students with sustained access to efficient and effective individualized instruction, may support improved civic outcomes for some SID in their post-school lives. An in-school emphasis on community experience (authentic inclusion in general education classrooms and other general settings) coupled with academic content instruction that promotes acquisition of skills and knowledge surrounding sociocultural standards and civic participation may...
help to improve post-school outcomes for some students with SID.

Social studies content includes civics, economics, geography and history. A December 2016 review indicated that 37 states require students to demonstrate proficiency in social studies or civics through assessment, and that every state includes civic learning or social studies in its academic standards or curriculum (Education Commission of the States, 2018). While federal testing mandates do not require states to administer summative assessment in social studies, it is important that educators, researchers, and teacher preparation programs focused on curriculum and instruction for students with SID consider state trends surrounding state curricula, standards, and summative assessment in the social studies. In addition to supporting post-school civic competency, educators teaching students with SID are responsible for providing grade-level state curriculum content instruction for their students and for preparing students to demonstrate academic proficiency through state alternate assessment (Every Student Succeeds Act, 2015; US Department of Education, 2015). To help educators fulfill these responsibilities, evidence based effective instructional practices for teaching social studies content for students with SID are needed.

A recognized challenge for educators serving students with SID is the provision of individualized instruction while maintaining authentic inclusive experiences within the flow of activities in general education classrooms (Jimenez & Kamei, 2015). Planning and implementation difficulties surrounding inclusive instruction can result in decreased group participation or diminished instructional rigor for students. For example, students may be removed from group activities within the general education classroom to participate in one-to-one instruction with a special education staff member. Embedded instruction (EI) has emerged as a strategy for promoting evidenced instruction for students with SID in general education classrooms (Kurth, Lyon, & Shogren, 2015). EI employs an instructional trial distribution schedule that coincides with the natural pauses and transitions that occur during general education instruction and has been used to teach academic content for students with SID. Jameson, McDonnell, Polychronis, and Riesen (2008) used constant time delay, peer-mediated instruction, and an EI distribution schedule to teach art definitions and health information described in middle school curriculum standards. Also, Collins, Evans, Creech-Galloway, Karl, and Miller (2007) used an EI distribution schedule to teach targeted math, science, and social studies words for students with SID in elementary, middle, and high school.

The method of instruction provided within an embedded instructional trial schedule may further help educators to promote authentic inclusion within the typical flow of events in the general education classroom. Some special education researchers have examined the use of structured inquiry-based instruction as an approximation of the instructional activities used by teachers and students in general education programs (Jimenez, Browder, Spooner, & Dibiase, 2012). Inquiry-based instruction involves teacher-guided question generation and examination of information resources to promote student learning instead of a direct teacher-to-student relay of facts. This method has been recognized as a way to promote student behaviors that support self-guided learning and information analysis in situations that do not involve direct oversight from others (National Institutes of Health, 2005). The National Council for the Social Studies recommends the use of an inquiry instruction process for teaching social studies concepts and skills. The process described by NCSS includes creation of a relevant question, selection of information resources to address the question, finding evidence in resources to support responses to the question, and sharing findings with others (National Council for the Social Studies [NCSS], 2017). As educators develop plans to provide social studies instruction for students with SID in general education settings, the use of a structured inquiry-based instructional framework may merit consideration. First, inquiry-based instruction is recommended to strengthen skills important to the promotion of civic competence and participation in democratic society for students (NCSS, 2017), an area in which students with SID need support (Agran & Hughes, 2013). Further, the framework allows an approximation of the activities and processes likely used by the general education...
class. Increasingly authentic inclusive experiences may occur as this approximation is shaped over time to more closely resemble general class and community-based activities.

Structured inquiry-based instructional frameworks or elements of these frameworks have been used to teach academic skills for students with SID. Jimenez, Browder, Spooner, and Dibiase (2012) used a structured inquiry-based instructional framework, peer-mediated embedded instruction, a graphic organizer, and constant time delay to teach science concepts to middle school students with moderate intellectual disability. Also, Wood, Browder, and Flynn (2015) taught middle school students with intellectual disability to generate and answer questions about social studies texts using a system of least prompts and a graphic organizer. Acquisition-level instruction in the latter study was conducted in self-contained special education classrooms.

To date, few studies inform instructional interventions within social studies content areas for students with significant intellectual disability. No studies exist that examine acquisition-level social studies content instruction for students with SID in general education settings using elements of inquiry-based instruction. The purpose of this study was to examine the effects of paraprofessional-mediated social studies instruction using structured inquiry-based instruction within an embedded instructional trial schedule on rates of correct responding to social studies concept questions for students with SID in junior high school. The research questions addressed in this study included (1) What are the effects of paraprofessional-delivered structured inquiry-based instruction, presented using an EI trial distribution schedule on rates of correct responding to social studies concept questions for students with SID in junior high school? (2) What are the effects of a behavioral skills training package on paraprofessionals’ ability to provide instruction using a structured inquiry-based instructional framework, embedded into the typical flow of activities in a general education classroom? (3) How do paraprofessionals and general education teachers rate the acceptability and effectiveness of the described intervention, with respect to intervention goals, procedures, and outcomes?

Method

Participants

Study participants included students and staff in a collaborating school district local to the University. Prior to the selection of study participants, all study procedures were reviewed and approved by University and school district Institutional Review Boards to ensure compliance with ethical standards for human subject research and adherence to local policies.

Student participants. Three junior high school students with intellectual disability were selected for study participation. Student participation criteria included (a) presence of a significant intellectual disability, as evidenced by qualification for and participation in the state alternate academic assessment, (b) adequate vision and hearing to interact with visually and orally presented instructional materials, (c) ability to make a choice among three or more visually presented pieces of text, pictures, or objects, (d) consistent participation in a general education social studies classroom with paraprofessional support, and (e) consistent attendance (two or fewer absences per month). Detailed student information is shown in Table 1.

Prior to the onset of study activities, the special education teacher helped to obtain signed parental consent for each student. The primary author met individually with students in the self-contained special education classroom to explain study activities and to gain oral and/or written assent.

Esteban. Esteban was a 13-year-old, eighth-grade student with a primary educational diagnosis of autism. He received educational services in a self-contained classroom for parts of the school day, and attended US history, science, dance, and art in general education classrooms with paraprofessional and peer support. During content instruction in the US history classroom, typical instructional supports included allowing him to copy peers’ written work and providing direct peer or paraprofessional support to complete activities.

Michaela. Michaela was an eighth-grade student who celebrated her 14th birthday during the course of the study. Her primary educational diagnosis was Other Health Impair-
ment (OHI). She received most of her educational services in a self-contained special education classroom and attended US history, science, and choir in general education classrooms with paraprofessional support to complete instructional activities. During content instruction in the US history classroom, Michaela typically traced words related to whole-class academic instruction or selected a leisure activity such as drawing or looking at picture books.

Scooter. Scooter was a 12-year-old, seventh-grade student with a primary educational diagnosis of intellectual disability. He received the majority of his daily academic instruction in a self-contained special education classroom and attended Utah studies and gym in general education settings with paraprofessional support. During content instruction in the Utah studies class, Scooter would complete worksheets adapted by the general education teacher, including coloring maps and drawing pictures of the concepts discussed during the class.

Paraprofessional participants. Two paraprofessionals implemented study activities with the participating students in the inclusive setting. Before beginning study activities, research staff obtained signed research participation consent from both paraprofessionals.

Robby. Robby was a 40-year-old white male who had two years of experience working as a special education paraprofessional at the junior high school. He had completed some college. Robby received informal and ongoing training at the junior high school concerning student-specific academic, behavioral, and other instructional strategies and interventions. He had completed yearly behavioral management and crisis de-escalation training and had received formal online-based training concerning autism spectrum disorders. Robby’s primary job at the school was to support students receiving services in the self-contained special education classroom.

Bea. Bea was a 55-year-old white female who had 10 years of experience working as a paraprofessional in special education. Bea had earned an associate’s degree in social sciences and was in the process of completing coursework towards a bachelor’s degree in psychology. She had participated in several formal and informal in-service trainings concerning disability, instruction of children with disabilities, and behavioral management and crisis de-escalation during her work as a paraprofessional in public schools.

General education teachers. Two teachers provided whole-class general education history instruction throughout the study. Prior to the onset of study activities, research staff obtained signed participation consent from both general education teachers. Mr. C taught eighth-grade US history and Ms. L taught seventh-grade Utah studies.

Settings

The study was conducted in a public junior high school in a mid-sized school district located in a large suburban community in the Mountain West region. The school served students through a traditional junior high school model, providing general and special education programs to neighborhood students in seventh through ninth grades.

### TABLE 1

Student Characteristics

<table>
<thead>
<tr>
<th>Student</th>
<th>Age</th>
<th>Grade Level</th>
<th>Gender</th>
<th>Race/Ethnicity</th>
<th>Primary Educational Diagnosis</th>
<th>Cognitive/Behavioral Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esteban</td>
<td>13yo</td>
<td>8</td>
<td>M</td>
<td>Mexican-American</td>
<td>ASD</td>
<td>Vineland-3: 62c</td>
</tr>
<tr>
<td>Michaela</td>
<td>14yo</td>
<td>8</td>
<td>F</td>
<td>White</td>
<td>OHI</td>
<td>ABAS-2: 50d</td>
</tr>
<tr>
<td>Scooter</td>
<td>12yo</td>
<td>7</td>
<td>M</td>
<td>White</td>
<td>ID</td>
<td>WISC-V: 49</td>
</tr>
</tbody>
</table>

the course of typical instructional activities in general education history classrooms. Michaela and Esteban attended US history in Mr. C’s classroom during a mid-morning period for four days each week, according to the general school schedule. Typical whole-group instructional activities in Mr. C’s class during the course of the study included watching movies depicting historical events, teacher-led lecture and discussion, textbook readings, and quiz reviews. Scooter attended Ms. L’s Utah studies class during an early morning period for four days each week, according to the general school schedule. Typical instructional activities in Ms. L’s class included whole-group teacher-led multimedia lessons (videos, slides), group and individual written work.

Probes. Probes to assess students’ rates of correct responding to social studies concept questions during instructional and maintenance phases were conducted during non-instructional times in the self-contained special education classroom.

Research Design

A multiple probe across participants design was used to establish experimental control and to inform the determination of a functional relationship between the intervention and student performance (Johnston, 2011). This design allowed for research staff to present instruction on each lesson sequentially. The design further allowed research staff to conduct intermittent rather than continuous assessment probes during non-instructional phases, potentially decreasing the impact of repeated exposures to the instructional materials on student responses (Horner & Baer, 1978).

Intervention Materials

The study authors created all intervention materials following consultation with special and general education teachers, a review of educational records describing student support needs and present levels of performance, and observations of students’ instruction in the special and general education classrooms. Materials were developed using Microsoft Word. Before study materials were presented to students during instruction and assessment, the general education teachers reviewed all materials for content validity and quality. They did not recommend any changes. Figure 1 shows a template of the instructional materials provided for Michaela and Scooter. Materials for Esteban were similar but excluded pictures and color framing.

Intervention materials were organized into four themed lessons for each of Utah studies and US history. Three of these lessons were used during intervention phases, and one lesson was used to measure generalization of the structured inquiry-based instructional process. Each themed lesson included three instructional texts aligned with major lesson concepts. Instructional texts were developed using class textbooks provided by general education teachers.

A graphic organizer and concept question cards were used during instruction to support students in the selection of a question relevant to the instructional text, as shown in Figure 1. Two response icons showed information presented in each instructional text. For Michaela and Scooter, these response icons included a picture and simplified wording to match sentences from the instructional text. For Esteban, response icons included highlighted words shown in the instructional text. Several distracter icons were created for use during instruction. As shown in Figure 1, each instructional text had an associated student response card containing two docking areas for response icons.

Paraprofessional Training and Procedural Fidelity

Paraprofessional training. Before paraprofessionals implemented the intervention with students, research staff provided small group training that included (a) information about recognizing when to embed instruction in the general education classroom, and (b) information about and practice using the inquiry-based social studies intervention. Information about embedding instruction was provided in a discussion format. Research staff verbally described embedded instruction, then provided several examples of appropriate times to embed the instructional intervention during general education social studies classes.

The training to prepare paraprofessionals to use the instructional intervention used a
behavioral skills training format (Ward-Horner & Sturmey, 2012), and was based on an instruction task analysis such as the example shown in Table 2. Intervention fidelity scores were calculated using this task analysis by determining a percentage: the number of correctly performed steps was divided by the total number of applicable steps for a role play performance, and then this number was multiplied by 100. Training concluded for a paraprofessional with achievement of an intervention fidelity score of 90% or greater during two consecutive role play sessions. Paraprofessional training again occurred prior to generalization instruction for each student, based on a generalization task analysis. Training methods and performance criterion were the same as previously described.

Procedural fidelity. Procedural fidelity was measured by calculating the percentage of task analysis steps correctly completed by paraprofessionals during instructional and general-
**TABLE 2**

**Instructional Task Analysis for Michaela and Scooter**

<table>
<thead>
<tr>
<th>Gains Student’s Attention, Presents Instructional Text, Presents Question Cards</th>
<th>Y</th>
<th>N</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Begins instruction during a natural pause in class instruction (transition, group work, individual work).</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>2. Presents the correct folder to the correct student.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>3. Gains student’s attention and brings student’s attention to folder.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>4. Reads each sentence in the instructional text, pointing to each picture.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>5. Places correct question card and two distracter question cards (in random order) along the left side of the folder.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>6. Points to folder title and says “What do we want to know?”</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>7. Points to each question card choice and reads the title.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>8. Says, “Pick one” (or similar) and waits 3 seconds.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>9. A. If student picks the correct card, indicates where student should place the card on the folder and provides verbal praise.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
<tr>
<td>9. B. If student picks the incorrect card or does not respond, pleasantly gets card from student, puts cards back on the table in a different order. Says, “What do we want to know? Pick one,” then uses a system of least prompts to indicate the correct answer and to indicate where student should place card on folder. Provides verbal praise.</td>
<td>Y</td>
<td>N</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Presents Response Icons**

<table>
<thead>
<tr>
<th>Y</th>
<th>N</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Places two correct and two distracter response icons on the question card.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2. Reads the question on the question card, then points to each response icon as the text is read on each one.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>3. Says, “Match one” (or similar) and waits 3 seconds.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4. A. If student correctly places an icon on top of a picture on the folder, provides verbal praise.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4. B. If student picks an incorrect icon or does not choose an icon, pleasantly gets icon from student, puts the icons back on the question card in a different order. Says, “Match one,” then uses a system of least prompts to indicate the correct answer and to indicate where student should place response icon on the folder. Provides verbal praise when student places the icon on the picture.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>5. Says, “Match another one” (or similar) and waits 3 seconds.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>6. A. Uses procedure described in 4a if student correctly places an icon.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>6. B. Uses procedure described in 4b if student incorrectly places an icon.</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

**Prompts Student to Place Response Icons on the Response Card**

<table>
<thead>
<tr>
<th>Y</th>
<th>N</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Points to title of the answer card and reads title of response card (ex. “Two facts about...”).</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2. Indicates to student to place each correct icon on the response card.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>3. Repeats title of the response card, points to each correct icon and reads text.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4. Provides enthusiastic social praise or a preferred tangible item.</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

**Resets Folder and Re-Presents Information**

<table>
<thead>
<tr>
<th>Y</th>
<th>N</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Removes all icons and cards.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2. Stores icons and cards in the Ziploc bag on the back of the folder.</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>3. Presents each of three concept statements two times during the class, in a rotating way – ex. Concept 1 – Concept 2 – Concept 3 – Concept 1 – Concept 2 – Concept 3. (Presents instruction to student 6 times -- pause at least 2-5 minutes before presenting another folder.)</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

| Date: | Staff: | Observer: | Percent marked Y: |
alization procedures in general education classrooms. This percentage was determined using the instructional or generalization task analyses for a student, and by using the same calculation procedures used during paraprofessional training.

The primary author observed paraprofessionals using the instructional or generalization procedures in the general education classroom for each student and simultaneously coded performance on each step using the appropriate task analysis. Corrective and reinforcing feedback was provided to paraprofessionals at the time of observations.

Dependent Variable and Measurement

Three dependent variables were assessed during the study. The primary dependent variable was the percentage of concept statements that students answered correctly per lesson. This variable was measured throughout a series of phases, including a baseline phase and three instructional phases. An additional phase was included to determine effects on this variable when students were presented with untaught instructional materials (generalization) and following lapses of time without additional instruction (maintenance).

Secondary dependent variables included the accuracy of paraprofessionals’ use of the instructional intervention in the general education setting following training, and paraprofessionals’ and general education teachers’ perceptions of factors supporting the social validity of the intervention.

Baseline. During the baseline phase, study authors sat side-by-side with student participants at a table in the self-contained special education classroom. Students did not have access to folders including instructional texts. To determine the rate of a student’s correct responses during a probe session, the first author presented response statement starter and a concept question card with two correct and two distracter response icons attached. She then read the lead-in statement shown on the concept statement starter card. For Michaela and Scooter, she also pointed to and read text on each of the response icons affixed to the concept question card. Error correction was not provided for incorrect responses, and students received verbal praise for engaging with the author/s and study materials.

Student responses to each probe were coded yes, no, or error. The author/s recorded yes if the student selected and then placed both of the correct response icons on the response starter statement card. No was recorded if the student placed zero or one correct icon on the starter statement card. Error was recorded if the probe trial was stopped and re-presented due to staff error, a student issue such as refusal or distraction, or other procedural error.

For each student, three baseline probes were conducted to determine rates of correct responding for three instructional lessons and one generalization lesson prior to the onset of instructional phases. Final baseline probes for each of these lessons were conducted just prior to the introduction of the first instructional phase.

Instructional phases. Following baseline, instructional phases occurred for each of the numbered lessons. An instructional phase change occurred for each student following mastery of a lesson. Mastery criterion for each lesson included (a) at least three days of instruction and concept statement assessment probes were conducted for the lesson, and (b) the student correctly answered 100% of the concept statement probes for at least two consecutive days per lesson.

During instructional phases, teaching occurred during the typical flow of activities in general education history classrooms. Each of three instructional texts within a lesson were presented twice during a US history class. During instruction, a paraprofessional presented instructional prompts for a student within the typical seating arrangement of the class and according to an instructional task analysis, such as the one shown in Table 2.

A system of least prompts with constant time delay procedure was used to teach students to use the instructional materials and to provide error correction. Paraprofessionals provided verbal praise for correct student responses to each step of the instructional task analysis.

Instructional probes. Instructional probes were used measure the dependent variable on each day a student participated in intervention during an instructional phase. During instructional probes, students sat side-by-side
with research staff at a table in the self-contained special education classroom. Procedures for presenting instructional probes and collecting data mirrored procedures used in the baseline phase. Error correction was not provided for incorrect responses, and students received verbal praise for interacting with researchers and study materials.

**Generalization.** For each student, a generalization lesson remained untaught throughout instructional phases. Students did not receive instruction for the three concepts within the generalization lesson or have access to instructional texts. The three concepts within the generalization lesson were probed on each day that instruction occurred.

Following the conclusion of the instructional phases, paraprofessionals embedded a generalization procedure during the course of general education activities. This was the first time that students were exposed to the instructional text and associated instructional materials for the generalization lesson. Paraprofessionals presented the instructional text to students and read each sentence in the text, pointing to pictures for Michaela and Scooter. They then placed the correct concept question card and two distracter question cards, each with four response icons attached, on the desk near the student. They also placed the student response card on the desk near the student and said, “It’s time to work.” Instruction was not provided as described in the instructional task analysis. Error correction was not provided for incorrect responses, and verbal praise was given for engaging with study materials and remaining on-task.

**Maintenance.** Maintenance probes occurred once weekly following the conclusion of all instructional phases and at least one week following mastery of a lesson. Maintenance probe procedures mirrored baseline probe procedures.

**Social validity.** Paraprofessionals and general education teachers reported their perceptions of intervention goals, processes, and outcomes using a Likert-type post-intervention assessment adapted from the Treatment Acceptability Rating Form used by Reimers and Wacker (1988). Paraprofessionals completed an assessment for each student they supported during intervention and each teacher completed one assessment to report perceptions for all student participants included in their classes. Table 3 shows the questions used to assess social validity and the associated staff responses.

**Results**

**Baseline and Instructional Probes**

The percentage of correct student responses to concept assessment probes throughout each phase of the study is shown in Figure 2. Three baseline probes were conducted for each of the numbered lessons and the generalization lesson for each participant. The percent of correct responses to all probes for all students during baseline remained at or below 33.3%. Following the onset of Lesson 1 instruction, rates of correct responding for Lesson 1 probes for each student showed an immediate and strong change in level and trend. All students achieved mastery criteria for Lesson 1 in three to four days of instruction.

Probes were again conducted for Lesson 2 just prior to instruction for this lesson to determine if learning had occurred since baseline. After instruction began for Lesson 2, rates of correct responding for all students showed a significant change in level and trend. All students achieved mastery criteria for Lesson 2 within three or four days of instruction.

Probes were conducted for Lesson 3 concepts just before the onset of instruction for this lesson. After Lesson 3 instruction began, correct responding for all students showed an increase in rate and trend. Students met mastery criteria for Lesson 3 within three days of instruction.

**Generalization**

Probes for the generalization lesson were conducted throughout baseline and instructional phases. Correct responding during probes for the generalization lesson during these phases remained low for all students. Following the onset of generalization instruction, there was an immediate positive change in the level of correct responding for concepts included in the generalization lesson for all students. Michaela met mastery criteria in seven instructional days, Scooter met mastery criteria in four instructional days, and Esteban met mastery criteria in three instructional days.
<table>
<thead>
<tr>
<th>Teacher Questions</th>
<th>Paraprofessional Questions</th>
<th>Response Options: Lower and Upper Values</th>
<th>General Education Teachers (N = 2, S = 2)</th>
<th>Paraprofessionals (N = 2, S = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How clear is your understanding of the procedures used during the instructional intervention with the target student/s?</td>
<td>1. How clear is your understanding of the procedures used during the instructional intervention with the target student?</td>
<td>Not at all clear 1  Very clear 6</td>
<td>Mean: 5  Range: NA</td>
<td>Mean: 6  Range: NA</td>
</tr>
<tr>
<td>2. How acceptable do you find the instructional intervention to be regarding what you know about how to support the target student/s in your class?</td>
<td>2. How acceptable do you find the instructional intervention to be regarding what you know about how to support the target student/s in specific history class?</td>
<td>Not at all acceptable 1  Very acceptable 6</td>
<td>Mean: 6  Range: NA</td>
<td>Mean: 5  Range: NA</td>
</tr>
<tr>
<td>3. How willing are you to promote the implementation the instructional intervention over time with the target student/s?</td>
<td>3. How willing are you to implement the instructional intervention over time with the target student?</td>
<td>Not at all willing 1  Very willing 6</td>
<td>Mean: 6  Range: NA</td>
<td>Mean: 5.3  Range: 5–6</td>
</tr>
<tr>
<td>4. Given the target student/s' academic support needs in your classroom, how reasonable do you find the instructional intervention?</td>
<td>4. Given the target student’s academic support needs in the specific history classroom, how reasonable do you find the instructional intervention?</td>
<td>Not at all reasonable 1  Very reasonable 6</td>
<td>Mean: 6  Range: NA</td>
<td>Mean: 4.7  Range: 4–6</td>
</tr>
<tr>
<td>5. How costly will it be to implement this instructional intervention (think about financial and material resources, as well as planning and materials preparation)?</td>
<td>5. How costly will it be to implement this instructional intervention (think about financial and material resources, as well as planning and materials preparation)?</td>
<td>Not at all costly 1  Very costly 6</td>
<td>Mean: 5.5  Range: 5–6</td>
<td>Mean: 4.7  Range: 4–6</td>
</tr>
<tr>
<td>6. How disruptive will it be to the usual flow of activities in your classroom to implement the instructional intervention?</td>
<td>6. How disruptive will it be to the usual flow of activities in the specific history classroom to implement the instructional intervention?</td>
<td>Not at all disruptive 1  Very disruptive 6</td>
<td>Mean: 3  Range: 2–4</td>
<td>Mean: 3  Range: 2–4</td>
</tr>
<tr>
<td>7. How affordable is this instructional intervention (consider money, materials, and time)?</td>
<td>7. How affordable is this instructional intervention (consider money, materials, and time)?</td>
<td>Not at all affordable 1  Very affordable 6</td>
<td>Mean: 5*  Range: NA</td>
<td>Mean: 3  Range: NA</td>
</tr>
</tbody>
</table>

(Continued on next page)
### Table 3 (Continued)

<table>
<thead>
<tr>
<th>Teacher Questions</th>
<th>Paraprofessional Questions</th>
<th>Response Options: Lower and Upper Values</th>
<th>General Education Teachers (N = 2, S = 2)</th>
<th>Paraprofessionals (N = 2, S = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. How much do you like the instructional intervention?</td>
<td>8. How much do you like the instructional intervention?</td>
<td>Do not like it at all</td>
<td>Like it very much</td>
<td>6</td>
</tr>
<tr>
<td>9. How much discomfort are the target students or general education students likely to experience as a result of this instructional intervention?</td>
<td>8. How much discomfort is the target student likely to experience as a result of this instructional intervention?</td>
<td>No discomfort at all</td>
<td>Very much discomfort</td>
<td>6</td>
</tr>
<tr>
<td>10. How willing would you be to make changes to usual classroom activities to promote the use of this instructional intervention for students with significant intellectual disability?</td>
<td>10. How willing would you be to make changes to the usual activities you do when supporting the target student in the specific history classroom to implement this instructional intervention?</td>
<td>Not at all willing</td>
<td>Very willing</td>
<td>6</td>
</tr>
<tr>
<td>11. How well will paraprofessionals’ use of this instructional intervention fit into your classroom routine?</td>
<td>11. How well will carrying out this instructional intervention fit into the specific history classroom routine?</td>
<td>Not at all well</td>
<td>Very well</td>
<td>6</td>
</tr>
</tbody>
</table>

Note. *One teacher did not answer this question. N = number of staff members completing assessments; S = number of assessments completed; NA = not applicable. Paraprofessionals completed an assessment for each of the student participants they supported during the study.*
Maintenance

Figure 3 shows the percentage of students’ correct responses to concept probes following lapses of time without additional instruction or exposure to the study materials. Scooter and Esteban maintained high rates of correct responding several weeks after instruction was discontinued for each lesson. Michaela’s rate of correct responding dropped to baseline levels after instruction was discontinued.

Social Validity

School staff responses to the social validity assessment are shown in Table 3. School staff generally reported that they found the inter-
vention to be acceptable and effective. Both the general education teachers and the para-professionals indicated that they liked the intervention (overall mean 5.2, with a range of 5–6), found it acceptable when considering students’ academic support needs in history (overall mean 5.4, with a range of 5–6), and would support the ongoing use of the intervention (overall mean 5.6, with a range of 5–6). One exception to the overall positive perception included the indication that the intervention was costly with respect to time and/or money (overall mean 5.0, with a range of 4–6), a finding often cited within the context of educational programs.

Figure 3. Percentage of correct responses following lapses of time after intervention for Michaela, Scooter, and Esteban (maintenance measures). Mastery L1 = day that mastery criteria for Lesson 1 was met; Mastery L2 = day that mastery criteria for Lesson 2 was met; Mastery L3 = day that mastery criteria for L3 was met; Mastery Gen = day that mastery criteria for the Generalization lesson was met.
interventions using author-developed instructional materials. Another exception included the indication that the intervention was somewhat disruptive to the flow of activities in general education classrooms (overall mean 3.0, with a range of 2–4).

Interobserver Agreement and Procedural Fidelity

Interobserver agreement. To collect interobserver data during concept assessment probes, the first author presented materials and prompts to students while another author observed. Each author independently recorded students’ responses to each probe. Interobserver agreement (IOA) was determined using a trial-by-trial calculation: the percentage of agreement was calculated by dividing the number of trial agreements by the total number of trials collected and then multiplying by 100. Interobserver data were collected for 76.7% of total probe sessions (within-student, within-phase range of 33.3–100%). IOA was 99.8% for all probes observed (within-student, within-phase range of 98.3–100%).

Procedural fidelity. The first author observed paraprofessionals and coded task analyses for 30% of all instructional sessions and 30.8% of all generalization sessions. The mean percentage of task analysis steps correctly completed during instructional sessions was 94.7%, with a range of 81–100%. The mean percentage of task analysis steps completed correctly by paraprofessionals during generalization sessions was 93.0%, with a range of 90.5–95.5%.

Discussion

The results of this study demonstrate the effectiveness of embedded structured inquiry-based instruction to teach social studies concepts for students with SID. Further, study results suggest that student participants used the structured inquiry-based instruction to acquire and recall information without prompting from others following repeated use of the process. These findings complement the literature supporting embedded instruction and structured inquiry-based instruction for students with SID and support the use of structured inquiry-based instruction in general education classrooms to teach social studies content for some students. Study findings also support the use of task analysis, time delay, and a system of least prompts to provide social studies instruction for students with SID (Wood et al., 2015), and provide evidence that these procedures can be used effectively in inclusive settings.

Limitations

There are limitations that merit consideration when interpreting the outcomes and implications of this study. First, the small number of study participants limits the generality of findings. While sufficient within-study replications occurred to demonstrate the effectiveness of embedded structured inquiry-based instruction for the particular study participants, it cannot be assumed that similar effects would occur for other participants. Further studies are needed before statements can be made about the effects of the interventions described in this study on a broader population of students and school staff.

Next, the academic content presented during study phases did not always synchronize with whole-class content. Although efforts were made to keep the content presented during the intervention temporally close to the general academic content of the classroom, there were several days when study participants received instruction on different concepts than the general body of the class. Due to this, it is possible that there was some loss of authenticity of the inclusive instruction provided during this study. This study did not use ongoing interdisciplinary (special education and general education) planning to determine lesson content for students with SID from day to day, but instead relied on materials that were prepared in advance and on a prescribed schedule of lesson presentation. The only other study using embedded instruction and an inquiry-based instructional framework to teach academic content also described this issue: to ensure that study participants kept pace with the changing content of a general education science class, Jimenez, Browder, Spooner, and Dibiase (2012) combined embedded instruction with 1:1 instruction with a special education teacher for three of five student participants. Future studies might examine schedules of content pre-
sentation that do not rely on the sequential mastery of lessons, but instead rely on the daily content presented in general education classrooms.

**Future Research**

Research is needed to replicate the effectiveness of embedded structured inquiry-based instruction to teach social studies content in inclusive settings. While some studies demonstrate student learning using elements of inquiry-based instruction to teach social studies (Schenning, Knight, & Spooner, 2012; Wood et al., 2015), future studies might provide further information about the efficacy of these procedures in general education classrooms. Additionally, future studies might examine the effectiveness of peer-mediated inquiry-based instruction in general education settings.

Also, research is needed to establish a link between social studies content instruction and post-school civic functioning for students with SID (for example, community access and participation, employment, and the exercising of societal rights such as voting). A first step may be developing measurable indicators of “civic competence” for students with SID. Currently, longitudinal information describing the correlation between any in-school curricular experiences and any post-school outcomes for students with SID is sparse. This information is needed to inform the in-school curricular and instructional experiences provided to students with SID to best promote their meaningful participation as adult members of society.

Finally, research to understand the school and district factors and systems processes involved in the successful implementation of inclusive academic interventions for students with SID is needed. The social validity assessment provided information about perceptions school staff held about the intervention described in this study. However, no follow-up activities such as interviews or focus groups were used to more clearly define staff rationales for their responses or to identify the local resources and processes needed to support continued use of the intervention. Research using organizational behavior management tools such as behavioral systems analysis and behavioral process analysis (Deiner, McGee, & Miguel, 2009) may help to identify systems factors that inhibit and support the continued use of inclusive academic instructional strategies.

**Implications for Practice**

The findings of this study suggest several practical implications. First, the use of an embedded instructional trial schedule to present inquiry-based social studies content instruction may offer a way for educators to provide instruction across curricular areas for some students with SID. Educators must juggle several instructional responsibilities as they provide educational services for students with SID, such as promoting authentic inclusive educational experiences, implementing rigorous and evidenced academic content instruction aligned with the grade-level curriculum, and providing instruction for the more functional and process-based skills that may lead to improved post-school civic outcomes for students. Post-school outcomes for students with SID may rely on educators’ capacities to simultaneously promote instruction across an array of academic and functional skills areas. Interventions such as the one described in this study may help to strengthen these educator capacities.

Next, instruction using a structured inquiry framework could provide a basis for the shaping of more complex questioning, researching, and communication skills that students with SID may use to complete academic work and to increase general problem-solving capacity. The intervention described in this study is a simplified version of the inquiry process described by the NCSS. The utility of the simplified process in promoting improved civic outcomes for students with SID may best be conceptualized as a “starting point.” As students gain fluency with a simplified structured inquiry process, more complex skills may be included in the instructional task analysis to more closely approximate activities in the general education classroom. For example, students might generate questions about instructional texts without a selection of question cards or might search for potential responses in adapted texts or web content.

Finally, additional collaboration and planning may be needed to promote inclusive in-
struction using embedded instruction when students use behaviors that can disrupt the learning environment. The social validity assessment showed that Scooter’s general education teacher and paraprofessional considered the instructional procedure to be mildly disruptive in the Utah Studies classroom. This finding was not supported by social validity assessment responses given by school staff supporting the other students. Scooter would sometimes verbally protest before beginning study activities in the history classroom or he would talk about study materials while working. These were common behaviors for Scooter in multiple instructional situations, and not exclusive to instruction provided during the study. The occurrence of disruptive behaviors warrants consideration within the context of embedded instruction: disruption is cited by school staff as a reason for the exclusion of students with SID from general education settings during academic instructional times. To reduce the potential for class interruptions for a student who uses disruptive behavior during instructional tasks, it may be necessary to pair EI with evidenced procedures to reduce the occurrence of disruptive behaviors.

Conclusion

This study provides evidence that some students with significant intellectual disability can learn social studies content in general education settings using embedded instruction and structured inquiry-based instruction. The instructional procedures described in this study may help to promote civic competence, authentic inclusive experiences, and access to grade-level general academic curriculum for some students with SID.

References


studies state standards. Retrieved from https://www.socialstudies.org/c3

Received: 20 September 2018
Initial Acceptance: 14 November 2018
Final Acceptance: 22 February 2019
AUTHORS

Aguayo, Virginia. See Hagiwara, Mayumi ........ 3
Alcalay, Aditt. Comparing multiple stimulus preference assessments without replacement to in-the-moment reinforcer analysis on rate of responding .............. 69
Amor, Antonio Manuel. See Hagiwara, Mayumi .......... 3
Antosh, Anthony A. See Burke, Kathryn M ...... 18
Athamanah, Lindsay S. Implementing a peer-mediated intervention in a work-based learning setting for students with autism spectrum disorders ......................... 196
Ault, Melinda J. See Chapman, Suzannah M ...... 161
Bahcalı, Turgut. Effectiveness of video modeling presented by tablet PC on teaching job interview skills to individuals with developmental disabilities ......... 249
Balint-Langel, Kinga. See Yuan, Chengan .......... 215
Barton, Erin E. See Ledford, Jennifer R ......... 315
Barton, Erin E. See Pokorski, Elizabeth A ..... 94
Bassette, Laura. Comparison of concrete and app-based manipulatives to teach subtraction skills to elementary students with autism ................. 391
Bottge, Brian A. See Chapman, Suzannah M .... 161
Bouck, Emily C. See Polloway, Edward A .... 30
Bouck, Emily. See Bassette, Laura ............ 391
Bowman, Jessica A. See Ryan, Joanna ......... 420
Brady, Michael P. See Kearney, Kelly B ........ 328
Burke, Kathryn M. Implementing evidence-based practices to promote self-determination: Lessons learned from a state-wide implementation of the self-determined learning model of instruction ....................... 18
Burke, Kathryn M. See Hagiwara, Mayumi .... 3
Centone, Kayla. Increasing peer-directed vocal manding of young children with autism spectrum disorder ..................... 377
Chapman, Suzannah M. Teaching algebra with a functional application to students with moderate intellectual disability .......... 161
Cheek, Aftynne E. Online module plus ecoaching: The effects on special education teachers’ comprehension instruction for students with significant intellectual disability .... 343
Chen, Hsinyi. See Li, Yi-Fan ..................... 263
Cihon, Joseph H. See Alcalay, Aditt .......... 69
Cihon, Joseph H. See Leaf, Justin B .......... 107
Cihon, Joseph H. See Leaf, Justin B ............ 274
Clark, Kelly A. Teaching soft skills to students with disabilities using UPGRADE Your Performance ..................... 41
Coleman, Olivia Fudge. See Ryan, Joanna .... 420
Conradi, Lyndsey Aiono. See Ryan, Joanna ... 420
Cremeraus, McKenzie. See Bassette, Laura .... 391
Cushing, Lisa S. See Athamanah, Lindsay S .... 196
Demchak, MaryAnn. Teachers’ Perception of Use and Actual Use of a Data-Based Decision-Making Process ............ 175
Dilks, Abigail. See Centone, Kayla .......... 377
Domire, Sarah. See Wolfe, Pamela ........... 186
Downey, Angelica. See Kearney, Kelly B .... 328
Dukes, Charles. See Kearney, Kelly B ....... 328
Eichelberger, Carrie. See Ryan, Joanna ....... 420
Ferguson, Julia L. See Alcalay, Aditt .......... 69
Ferguson, Julia L. See Leaf, Justin B .......... 274
Fujiiwara, Aya. Promoting social play based on ecological assessment and social play selection conditions of a child with autism spectrum disorder in an inclusive early childhood classroom ....................... 288
Gage, Nicholas A. See Luo, Li ............... 358
Gallup, Jennifer. See Kocaoz, Onur Emre ...... 237
Gardner, Stephanie. See Wolfe, Pamela ..... 186
Gilson, Carly B. See Li, Yi-Fan .............. 263
Goo, Minokowan. Systematic review of using portable smart devices to teach functional skills to students with intellectual disability .. 57
Hagiwara, Mayumi. International trends in inclusive education intervention research: A literature review ...................... 3
Hu, Xiaoyi. Using an individual work system to increase independence for students with autism in a special education classroom in China ............... 119
Hua, Youjia. See Yuan, Chengan ............. 215
Jameson, J. Matt. See Ryan, Joanna .......... 420
Jimenez, Bree A. See Cheek, Aftynne E .... 343
Johnston, Susan S. See Ryan, Joanna ........ 420
Kaldenberg, Erica. See Saletta, Meredith .... 225
Kearney, Kelly B. Push hard and fast: Teaching college students with intellectual disability to perform hands-only cardiopulmonary resuscitation ........................................ 328
Kim, Soyeon. See Roberts, Carly A .............. 147
Kocaoz, Onur Emre. Impact of video modeling combined with skillstreaming teaching procedures on the social interaction skills of middle school-aged children with ASD ........ 237
Konrad, Moira. See Clark, Kelly A ............... 41
Kucharczyk, Suzanne. Parent implementation of structured work systems on child acquisition of independence skills through family preferred routines .............................. 83
LaPlante, Terri See Burke, Kathryn M ........... 18
Leaf, Justin B. Comparing no-no prompt to flexible prompt fading to teach expressive labels to individuals diagnosed with autism spectrum disorder .................................... 274
Leaf, Justin B. Descriptive analysis of the use of punishment-based techniques with children diagnosed with autism spectrum disorder ........................................ 107
Leaf, Justin B. See Alcalay, Aditt ................. 69
Leaf, Ronald. See Alcalay, Aditt ................. 69
Leaf, Ronald. See Leaf, Justin B ................. 107
Leaf, Ronald. See Leaf, Justin B ................. 274
Ledford, Jennifer R. See Pokorski, Elizabeth A . 94
Ledford, Jennifer R. Visual display of graphic data in single case design studies: Systematic review and expert preference analysis .................................................. 315
Lee, Gabrielle T. See Hu, Xiaoyi ................. 119
Li, Yi-Fan. Effects of a self-monitoring strategy to increase classroom task completion for high school students with moderate intellectual disability ................................... 263
Little, Mary E. See Kocaoz, Onur Emre ........ 237
Liu, Yanhong. See Hu, Xiaoyi ................. 119
Luo, Li. Systematic review of functional communication training in early care and education settings ...................................................... 358
Maffini, Cara S. See Simpson, Lisa A .......... 301
Maurer, Adela L. See Goo, Minokowan ....... 57
McDonnell, John. See Ryan, Joanna ............ 420
McEachin, John See Leaf, Justin B .......... 274
McEachin, John. See Alcalay, Aditt .......... 69
McEachin, John. See Leaf, Justin B .......... 107
Meyer, Nancy. See Roberts, Carly A .......... 147
Mitchell, Erin. See Leaf, Justin B .......... 107
Mrla, Tiffany. See Kucharczyk, Suzanne ....... 83
Olcay-Gul, Seray. Effectiveness of teaching social skills to individuals with autism spectrum using cool versus not cool ............................................ 132
Ozen, Arzu. See Bahcali, Turgut ................. 249
Park, Jiyoung. See Bassette, Laura .............. 391
Pokorski Elizabeth A. See Ledford, Jennifer R ... 315
Pokorski, Elizabeth A. Comparison of antecedent activities for increasing engagement in a preschool child with ASD during a small group activity ........................................... 94
Polloway, Edward A. Educational programs for students with intellectual disability: Demo- graphic patterns ............................................................. 30
Prykanowski, Debra A. See Luo, Li ............... 358
Raley, Sheida K. See Burke, Kathryn M ........ 18
Rivera, Kiara. See Saletta, Meredith ............. 225
Roberts, Carly A. Using content area literacy strategies during shared reading to increase comprehension of high school students with moderate intellectual disability on adapted science text ........................................... 147
Rock, Marcia L. See Cheek, Afyynne E ......... 343
Rudrud, Eric H. See Alcalay, Aditt ............... 69
Ruiz, Salvador. See Wolfe, Pamela ............... 186
Ryan, Joanna. Inclusive inquiry-based social studies instruction for students with significant intellectual disability .................................................... 420
Sadler, Kate M. Video self-modeling and functional behavior assessment to modify aggressive behaviors in students with autism spectrum disorder and intellectual disabili- ties ................................................................. 406
Saletta, Meredith. Do illustrations promote reading comprehension in adults with intellectual or developmental disabilities? ........ 225
Schuck, Rachel K. See Simpson, Lisa A .......... 301
Schulze, Kimberly A. See Alcalay, Aditt ........ 69
Severini, Katherine E. See Ledford, Jennifer R ... 315
Shepley, Sally B. See Chapman, Suzannah M ... 161
Shogren, Karrie A. See Burke, Kathryn M ...... 18
Shogren, Karrie A. See Hagiwara, Mayumi ...... 3
Shurr, Jordan. See Bassette, Laura ............... 391
Simpson, Lisa A. Examining use of school personnel in CBT interventions for anxiety in students with ASD ................................................. 301
Sonoyama, Shigeki. See Fujiwara, Aya ......... 288
Spriggs, Amy D. See Chapman, Suzannah M ... 161
Sutter, Chevonne. See Demchak, MaryAnn ........ 175
Tandy, Jacob. See Roberts, Carly A .......... 147
Taubman, Mitchell. See Leaf, Justin B .......... 107
Taylor, Abigail L. See Pokorski, Elizabeth A ...... 94
Test, David W. See Clark, Kelly A ............... 41
Comparing multiple stimulus preference assessments without replacement to in-the-moment reinforcer analysis on rate of responding. Aditt Alcalay, Julia L. Ferguson, Joseph H. Chion, Norma Torres, Justin B. Leaf, Ronald Leaf, John McEachin, Kimberly A. Schulze, and Eric H. Rudrud

Comparing no-no prompt to flexible prompt fading to teach expressive labels to individuals diagnosed with autism spectrum disorder. Justin B. Leaf, Joseph H. Chion, Julia L. Ferguson, Ronald Leaf, and John McEachin

Comparison of antecedent activities for increasing engagement in a preschool child with ASD during a small group activity. Elizabeth A. Pokorski, Erin E. Barton, Jennifer R. Leaf, Ronald Leaf, Mitchell Taubman, and John McEachin

Comparison of concrete and app-based manipulatives to teach subtraction skills to elementary students with autism. Laura Bassette, Emily Bouch, Jordan Shurr, Jiyoung Park, and McKenzie Cremeans


Do illustrations promote reading comprehension in adults with intellectual or developmental disabilities? Meredith Saleetta, Erica Kaldenberg, Kiara Rivera, and Audrey Wood

Educational programs for students with intellectual disability: Demographic patterns. Edward A. Polloway, Emily C. Bouch, and Lihua Yang

Effectiveness of teaching social skills to individuals with autism spectrum using cool versus not cool. Seray Olcay-Gul and Sezgin Vuran

Effectiveness of video modeling presented by tablet PC on teaching job interview skills to individuals with developmental disabilities. Turgut Bahcalı and Arzu Ozen

Effects of a self-monitoring strategy to increase classroom task completion for high school students with moderate intellectual disability. Yi-Fan Li, Hsinyi Chen, Dalun Zhang, and Carly B. Gilson

Effects of constant time delay on route planning using Google maps for individuals with intellectual and developmental disabilities

Zimmerman, Kathleen N. See Ledford, Elizabeth A

Wolfe, Pamela. Review of instructional strategies used in socio-sexuality education for individuals

Wood, Audrey. See Saleetta, Meredith

Yang, Lihua. See Wolfe, Pamela

Yuan, Chengan. Effects of constant time delay on route planning using Google maps for young adults with intellectual and developmental disabilities

Zhang, Dalun. See Li, Yi-Fan

Zimmerman, Kathleen N. See Ledford, Jennifer R

TITLES

Comparing multiple stimulus preference assessments without replacement to in-the-moment reinforcer analysis on rate of responding. Aditt Alcalay, Julia L. Ferguson, Joseph H. Chion, Norma Torres, Justin B. Leaf, Ronald Leaf, John McEachin, Kimberly A. Schulze, and Eric H. Rudrud

Comparing no-no prompt to flexible prompt fading to teach expressive labels to individuals diagnosed with autism spectrum disorder. Justin B. Leaf, Joseph H. Chion, Julia L. Ferguson, Ronald Leaf, and John McEachin

Comparison of antecedent activities for increasing engagement in a preschool child with ASD during a small group activity. Elizabeth A. Pokorski, Erin E. Barton, Jennifer R. Ledford, Abby L. Taylor, Elisabeth Johnson, and Heather K. Winters

Comparison of concrete and app-based manipulatives to teach subtraction skills to elementary students with autism. Laura Bassette, Emily Bouch, Jordan Shurr, Jiyoung Park, and McKenzie Cremeans


Do illustrations promote reading comprehension in adults with intellectual or developmental disabilities? Meredith Saleetta, Erica Kaldenberg, Kiara Rivera, and Audrey Wood

Educational programs for students with intellectual disability: Demographic patterns. Edward A. Polloway, Emily C. Bouch, and Lihua Yang

Effectiveness of teaching social skills to individuals with autism spectrum using cool versus not cool. Seray Olcay-Gul and Sezgin Vuran

Effectiveness of video modeling presented by tablet PC on teaching job interview skills to individuals with developmental disabilities. Turgut Bahcalı and Arzu Ozen

Effects of a self-monitoring strategy to increase classroom task completion for high school students with moderate intellectual disability. Yi-Fan Li, Hsinyi Chen, Dalun Zhang, and Carly B. Gilson

Effects of constant time delay on route planning using Google maps for young adults with intellectual and developmental disabilities. Chengan Yuan, Kinga Balint-Langel, and Youjia Hua

Examining use of school personnel in CBT interventions for anxiety in students with ASD. Lisa A. Simpson, Cara S. Maffini, and Rachel K. Schuck

Impact of video modeling combined with skill-streaming teaching procedures on the social interaction skills of middle school-aged children with ASD. Onur Emre Kocaoz, Mary E. Little, and Jennifer Gallup

Implementing a peer-mediated intervention in a work-based learning setting for students with autism spectrum disorders. Lindsay S. Athamanah and Lisa S. Cushing

Implementing evidence-based practices to promote self-determination: Lessons learned from a state-wide implementation of the self-determined learning model of instruction. Kathryn M. Burke, Karrie A. Shogren, Sheila K. Raley, Michael L. Wehmeyer, Anthony A. Antosh, and Terri LaPlante

Inclusive inquiry-based social studies instruction for students with significant intellectual disability. Joanna Ryan, J. Matt Jameson, Olivia Fudge Coleman, Carrie Eichelberger, Jessica A. Bowman, Lyndsey Aiono Conrad, Susan S. Johnston, and John McConnell

Index–Volume 54, 2019 / 439
Increasing peer-directed vocal manding of young children with autism spectrum disorder. Kayla Centone, Matt Tincani, and Abigail Dilks .............................. 377

International trends in inclusive education intervention research: A literature review. Mayumi Hagiwara, Antonio Manuel Amor, Karrie A. Shogren, James R. Thompson, Miguel Angel Verduco, Kathryn M. Burke, Hatice Uyanik, and Virginia Aguayo ...... 3

Online module plus ecoaching: The effects on special education teachers’ comprehension instruction for students with significant intellectual disability. Aftynne E. Cheek, Marcia L. Rock, and Bree A. Jimenez .......... 343

Parent implementation of structured work systems on child acquisition of independence skills through family preferred routines. Suzanne Kucharczyk, Peggy J. Schaefer Whitby, and Tiffany Mrla .................. 83

Promoting social play based on ecological assessment and social play selection conditions of a child with autism spectrum disorder in an inclusive early childhood classroom. Aya Fujiwara and Shigeki Sonoyama ............................. 288

Push hard and fast: Teaching college students with intellectual disability to perform hands-only cardiopulmonary resuscitation. Kelly B. Kearney, Michael P. Brady, Charles Dukes, and Angelica Downey .................... 328


Systematic review of functional communication training in early care and education settings. Li Luo, Nicholas A. Gage, and Debra A. Prykanowski ................................. 358

Systematic review of using portable smart devices to teach functional skills to students with intellectual disability. Minokowan Goo, Adela L. Maurer, and Michael L. Wehmeyer ...... 57

Teachers’ Perception of Use and Actual Use of a Data-Based Decision-Making Process. Mary-Ann Demchak and Chevonne Sutter ............ 175

Teaching algebra with a functional application to students with moderate intellectual disability. Suzzannah M. Chapman, Melinda J. Ault, Amy D. Spriggs, Brian A. Bottge, and Sally B. Shepley ............................. 161

Teaching soft skills to students with disabilities using UPGRADE Your Performance. Kelly A. Clark, David W. Test, and Moira Konrad ................................. 41

Using an individual work system to increase independence for students with autism in a special education classroom in China. Xiaoyi Hu, Gabrielle T. Lee, Yanhong Liu, and Manman Wu ................................. 119

Using content area literacy strategies during shared reading to increase comprehension of high school students with moderate intellectual disability on adapted science text. Carly A. Roberts, Soyeon Kim, Jacob Tandy, and Nancy Meyer ................................. 147

Video self-modeling and functional behavior assessment to modify aggressive behaviors in students with autism spectrum disorder and intellectual disabilities. Kate M. Sadler ............... 406

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
Differentiating Instruction in the Inclusive Classroom

STRATEGIES FOR SUCCESS

BARBARA GARTIN
NICKI MURDICK
BORELLE PEPPER
MARCIA IMBERU
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/](http://pubs.cec.sped.org/p6296/)
Statement of Ownership, Management, and Circulation  
(Act of August 12, 1970, Section 3685 Title 39, United States Code)

1. Publication Title: EDUCATION AND TRAINING IN AUTISM AND DEVELOPMENTAL DISABILITIES.  
2. Publication Number: 0016-8500.  
3. Filing Date: September 25, 2019.  
4. Issue Frequency: Quarterly in March, June, September, and December.  
5. No. of Issues Published Annually: 4.  
6. Annual Subscription Price: $100 domestic; $104 foreign; $249 institution; $254 foreign.  
9. Full Names and Complete Mailing Addresses of Publisher and Editor: Publisher—The Council for Exceptional Children, 3100 Clarendon Blvd., Suite 600, Arlington, Virginia 22201-5332; Editor—Dr. Stanley H. Zucker, Farmer 402, Arizona State University, PO Box 871811, Tempe, AZ 85287-1811; Managing Editor—Dr. Stanley H. Zucker, Farmer 402, Arizona State University, PO Box 871811, Tempe, AZ 85287-1811.  
11. Known Bondholders, Mortgagees, and Others: Not applicable.  
12. Tax Status: The purpose, function, and nonprofit status of this organization and the exempt status for Federal Income Tax purposes have not changed during the preceding 12 months.  
13. Publication Name: EDUCATION AND TRAINING IN AUTISM AND DEVELOPMENTAL DISABILITIES.  
15. Extent and nature of circulation:  

<table>
<thead>
<tr>
<th></th>
<th>Average no. copies ea. issue for past 12 mos.</th>
<th>No. of copies of issue nearest filing date</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Total no. copies printed</td>
<td>2,237</td>
<td>2,194</td>
</tr>
<tr>
<td>B. Paid circulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Sales through dealers and carriers, street vendors, and counter sales</td>
<td>55</td>
<td>50</td>
</tr>
<tr>
<td>2. Mail subscription</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total paid circulation</td>
<td>2,086</td>
<td>2,044</td>
</tr>
<tr>
<td>D. Free distribution by mail; samples, complimentary, and other free copies</td>
<td>2,141</td>
<td>2,094</td>
</tr>
<tr>
<td>E. Free distribution outside the mail; carriers or other means</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>F. Total free distribution</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G. Total distribution</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>H. Copies not distributed</td>
<td>2,166</td>
<td>2,124</td>
</tr>
<tr>
<td>1. Office use, left-over, unaccounted, spoiled after printing</td>
<td>71</td>
<td>70</td>
</tr>
<tr>
<td>2. Returns from news agents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. Total</td>
<td>2,237</td>
<td>2,194</td>
</tr>
<tr>
<td>J. Percent paid and/or requested circulation</td>
<td>98.84%</td>
<td>98.58%</td>
</tr>
</tbody>
</table>

17. This statement of ownership will be printed in the VOL. 54 #4 (12/01/2019 issue of this publication.  
18. Signature and Title:  

Teresa Doughty Executive Director September 25, 2019
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.
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.
21st International Conference on Autism, Intellectual Disability & Developmental Disabilities

Research-Informed Practice

January 22 - 24, 2020
Sarasota, Florida

CEC’s Division on Autism and Developmental Disabilities (DADD) conference will have innovative and engaging presentations that explicitly link research to practice within the following or related topical areas:

- Academic Skills
- Assistive & Adaptive Tech
- Life Skills
- Sexuality
- Mental Health
- Post-secondary Initiatives
- Communication
- Social Skills
- Evidence-based Practices
- Self-determination
- Transition
- Assessment

For further information, please contact Cindy Perras, DADD’s Conference Co-ordinator: cindy.perras@gmail.com.