Education and Training in Autism and Developmental Disabilities

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

Volume 52  Number 3  September 2017
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  David Gihak  Michael Mayton  Keith Storey
Kevin Ayres  Ginevra Courtade  John McDonnell  William Therrien
Devender Banda  Samuel A. DiGangi  Hedda Meadan-Kaplansky  Matt Tincani
Juliet Hart Barnett  Teresa Doughy  Linda Mechling  Jason Travers
Laura Bassette  Kimberly W. Fisher  Pam Mims  Toni Van Laarhoven
Kyle D. Bennett  Carolyn Hughes  Wendy Oakes  Elizabeth West
Emily Bouck  Bree Jimenez  Karrie Shogren  John Wheeler
Amanda Boutot  Russ Lang  Scott Sparks  Mark Wolery
Michael P. Brady  Justin Leaf  Fred Spooner  Leah Wood
Stacy Carter  Wendy Machalicek  Robert Stoddten  Dalun 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, 2900 Crystal Drive, Suite 1000, Arlington, VA 22202-3557. Advertising rates are available upon request.

Manuscripts should be typed, double spaced, and sent (three copies) 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 cover sheet that gives the names, affiliations, and complete addresses of all authors.

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

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


Division on Autism and Developmental Disabilities
The Council for Exceptional Children

Board of Directors

Officers
PAST PRESIDENT David Gihak
PRESIDENT Elizabeth West
PRESIDENT-ELECT Jordan Shurr
VICE PRESIDENT Michael Wehmeyer
SECRETARY Meaghan McCollow
TREASURER Gardiner Umbarger

Members
Autumn Eyre (Student Representative)
Beth Kavannagh
Nikki Murdieck
Leah Wood

Executive Director
Teresa Doughty
Publications Chair
Michael Wehmeyer
Communications Chair
Emily C. Bouck
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, 2900 Crystal Drive, Suite 1000, Arlington, Virginia 22202-3557. 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, 2900 Crystal Drive, Suite 1000, Arlington, Virginia 22202-3557.
International Collaboration and Its Contributions: Disseminating Knowledge and Supporting Evidence-Based Practices Across Countries
BELVA C. COLLINS, ELIF TEKIN-IFTAR, and SERAY OLÇAY-GÜL

Early Intervention in Autism Spectrum Disorder: Concerns and Support of Portuguese Mothers
LUCIA S. MAGALHAES and ANA P. S. PEREIRA

Comprehension of Verbal and Visual Metaphors among Individuals with Intellectual Disability with and without Down Syndrome
SHLOMIT SHNITZER-MEIROVICH, HEFZIBA LIFSHITZ-VAHAB, and NIRA MASHAL

Functional Living Skills and Adolescents and Adults with Autism Spectrum Disorder: A Meta-Analysis
EE REA HONG, JENNIFER B. GANZ, KRISTI MORIN, JOHN L. DAVIS, JENNIFER NINCI, LESLIE NEELY, and MARGOT B. BOLES

Effects of Modeling, Story Templates, and Self-Graphing in the Use of Story Elements by Students with Moderate Intellectual Disability
ROBERT PENNINGTON and MALLORY KOEHLER

Effects of Video Modeling on the Instructional Efficiency of Simultaneous Prompting Among Preschoolers with Autism Spectrum Disorder
DERYA GENC-TOSUN and ONUR KURT

The Role of the Replacement Behavior in Function-Based Intervention
LINDA M. REEVES, JOHN UMBREIT, JOLENEA B. FERRO, and CARL J. LIAUPSIN

Concrete and App-Based Manipulatives to Support Students with Disabilities with Subtraction
EMILY C. BOUCK, COURTNEY CHAMBERLAIN, and JIYON PARK

An Analysis of State Guidelines for Intellectual Disability
EDWARD A. POLLOWAY, MARYLEEN AUGUSTE, J. DAVID SMITH, and DELIA PETERS

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

December 2017

Educational outcomes for secondary students with mild intellectual disability. Emily C. Bouck, College of Education, Michigan State University, 620 Farm Lane, 335 Erickson Hall, East Lansing, MI 48824.

Single case design elements in text comprehension research for students with developmental disabilities. Sara M. Snyder, Victoria F. Knight, Kevin M. Ayres, Pamela J. Mims, and Emily C. Sartini, Department of Communication Sciences and Special Education, University of Georgia, 570 Aderhold Hall, 110 Carlton Street, Athens, GA 30602.

Parents’ educational expectations for young children with autism spectrum disorder. Hillary H. Bush, Shana R. Cohen, Abbey S. Eisenhower, and Jan Blacher, Learning and Emotional Assessment Program (LEAP), Massachusetts General Hospital, Department of Psychiatry, 151 Merrimac Street, 5th Floor, Boston, MA 02114.

Parametric analysis of specific praise rates on the on-task behavior of elementary students with autism. Michael P. Kranak, Sheila Alber-Morgan, and Mary R. Sawyer, The Ohio State University, A356 PAES Bldg., 305 W. 17th Avenue, Columbus, OH 43210.

Effects of school staff communication on repair strategies of students with severe intellectual disability. Orit E. Hetzroni and Maayan Shalev, Department of Special Education, University of Haifa, Mount Carmel, Haifa 31905 ISRAEL.

Comparing self-reinforcement to teacher-delivered reinforcement during activity schedule completion on the iPod Touch. Brittany N. Beaver, Sharon A. Reeve, Kenneth F. Reeve, and Ruth M. DeBar, Department of Applied Behavior Analysis, Caldwell University, 120 Bloomfield Avenue, Caldwell, NJ 07006.

Efficiency of an electronic editing strategy with college students with intellectual and developmental disabilities. Suzanne Woods-Groves, Youjia Hua, Jeremy W. Ford, and Katelyn M. Neil, Dept. of Teaching and Learning, 246 Lindquist Center N, Iowa City, IA 52242-1529.

Inclusion of students with significant disabilities in SWPBS evaluation tools. Jennifer A. Kurth, Alison Zagona, Mayumi Hagiwara, and Matt Enyart, University of Kansas, 1122 W. Campus Road, Joseph R. Pearson Hall Room 541, Lawrence, KS 66045.

Exploring predictors of postsecondary outcomes for students with autism spectrum disorder. Amy Nasamran, Sara E. Witmer, and James E. Los, Michigan State University, 620 Farm Lane, Rm. 434, East Lansing, MI 48824.

Address is supplied for author in boldface type.
Abstract: This article explores how international collaboration among researchers can contribute to developing evidence-based practices and disseminating knowledge in the field of special education. A review of a sample of special education journals published in English to identify articles written in collaboration by researchers from different countries is followed by an example of a collaborative relationship between special education researchers in Turkey and the United States that illustrates the potential of international collaboration to impact evidence-based practices. The article concludes with a discussion of the importance of international collaboration and recommendations for others who are interested in international collaboration.

For the past few decades, a great emphasis has been placed on developing evidence-based practices and disseminating knowledge in the field of special education. The history of the efforts to develop evidence-based practices can be grouped into two milestones: (a) understanding the need for and developing the definition of evidence-based practices and (b) translating research into practice. These two milestones can be considered as different but mutually complementary, and neither has yet been thoroughly defined. As a bridge, evidence-based practices have a special role in closing the gap between research and practice. The Council for Exceptional Children (CEC) has created professional standards that are recognized internationally, and these include “the use of a repertoire of evidenced-based instructional strategies to advance learning of individuals with exceptionalities” (CEC, 2015, p. 3). Specifically, in the United States, national educational policies stemming from the No Child Left Behind Act (NCLB) and the Individuals with Disabilities Education Act (IDEA) require that special education teachers use evidence-based practices in their classrooms; however, many teachers of students with disabilities have implemented teaching practices that have shown little or no effect on student outcomes (Cook & Schirmer, 2003). Researchers continue to pay great attention to providing input to understanding and defining evidence-based practices, and policymakers and lawmakers have been working to create a framework for educators and practitioners to use them to better serve and maximize the outcomes of education for children with disabilities.

The purpose of scientific research in special education is to describe, predict, and control the phenomena in studies and communicate the results to consumers and policymakers to effect better services. Different types of research methods yield knowledge at different levels. A generally accepted tenet of educational research holds that research designs...
exhibiting experimental control yield the most appropriate/strong answers to questions regarding whether or not a practice works (Cook, Tankersley, Cook, & Landrum, 2008). Since the beginning of the new millennium, there have been efforts to define the quality indicators of research methods used in special education (Brantlinger, Jimenez, Klinger, Pugachi, & Richardson, 2005; Cook & Cook, 2011; Gersten et al., 2005; Horner et al., 2005; Odom & Strain, 2002; Odom et al., 2005; Thompson, Diamond, William, Snyder, & Snyder, 2005).

Single subject research studies have played an important role in the development of evidence-based practices and have come to be recognized as rigorous scientific methodology (Gast, 2010). In these designs, a functional relationship between the independent variable and the dependent variable is established through repeated demonstrations of effect, with design features (e.g., stable baseline data, time-lagged intervention, withdrawal and reinstatement of treatment) to control for threats to internal and external validity. Horner et al. (2005) described the criteria to be used in single subject research to identify evidence-based practices in special education. Single subject research documents a practice as evidence-based when

(a) the practice is operationally defined, (b) the context in which the practice is to be used is defined, (c) the practice is implemented with fidelity, (d) the results from a single subject research documents a functional relationship between dependent and independent variables, and (d) the experimental effects are replicated across a sufficient number of studies, researchers, and participants to allow confidence in the findings (p. 175–176).

They also recommended that, for an intervention to be regarded as evidence-based, there must be at least five high-quality single-subject research studies demonstrating experimental control that “were published in peer-reviewed journals, (b) were conducted in at least three different geographical locations, (c) were conducted by at least three different researchers, and (d) included a minimum of 20 participants across studies. Among these criteria, two of them (i.e., items b and c) can be supported by establishing international collaboration between researchers who study the same line of research in the field” (p. 176).

The specific goal of this article is to describe how researchers can influence and establish evidence-based practices through international collaboration. In this paper, specifically (a) review single subject studies written in collaboration by researchers from different countries in a sample of special education journals, (b) share an example of an international collaboration that has supported evidence-based practices through publications and presentations, and (c) discuss the impact of collaborative research and how it can be facilitated.

Articles Written in Collaboration by Researchers from Different Countries in Special Education Journals

We first conducted a review of the professional special education literature from 2010–2015 to identify research studies that were published through collaboration by researchers across countries. We limited the search to a sample of seven journals that publish research in the field of special education: Education and Training in Autism and Developmental Disabilities (ETADD), Exceptional Children (EC), International Journal of Special Needs Education (IJSNE), Teacher Education and Special Education (TESE), The Journal of Early Intervention (JEI), The Journal of Special Education (JSE), and Topics in Early Childhood Special Education (TOPICS). Although these are not inclusive of all journals, especially those that are dedicated to specific areas of special education (e.g., ABA, early childhood, technology, transition) or specific disabilities (e.g., autism spectrum disorder, behavior disorders, learning disabilities, severe disabilities), these journals represent those that are typically read by faculty from Turkey and are meant to be representative of journals published in English that might influence international research. We reviewed each issue of these journals to identify articles written via international collaboration based on the listed affiliations of the authors. Once we identified research articles written in these journals that appeared to be the product of international collaboration,
we selected those articles for further analysis that used single subject research methodology because this is a common design used in special education to identify evidence-based practices. These articles are marked with an * in the references.

Coding the Articles and Interrater Reliability

The consistency between two of the co-authors was 99% agreement for detecting the articles written through collaborative efforts in the sample of journals we reviewed. Our search resulted in 52 (.06%) journal articles that met the inclusion criteria out of 925 published articles. Table 1 summarizes the total number of peer-reviewed articles in these journals and the articles written in collaboration across authors in different countries between 2010–2015. Among the articles published through international collaboration, only 26.9% \( (n = 14) \) consisted of data-based research (either group or single subject), and 64.3% \( (n = 9) \) of these used single subject research methodology. We coded each of the single subject research studies selected for analysis using the following parameters: (a) journal in which article was published, (b) names of authors, (c) topics of study, (d) country of each author’s affiliation, (e) type of experimental design, and (f) funding received to conduct the study. The second and third authors coded these articles independently using these parameters and attained 100% agreement. The articles

<table>
<thead>
<tr>
<th>Journals</th>
<th>Number of Articles Published Between 2010–2015</th>
<th>Number and Percent of Articles Written in Collaboration</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>149</td>
<td>7/4.7</td>
</tr>
<tr>
<td>ETADD</td>
<td>261</td>
<td>13/4.9</td>
</tr>
<tr>
<td>IJSNE</td>
<td>42</td>
<td>11/26</td>
</tr>
<tr>
<td>JEI</td>
<td>103</td>
<td>5/0.05</td>
</tr>
<tr>
<td>JSE</td>
<td>118</td>
<td>8/0.07</td>
</tr>
<tr>
<td>TESE</td>
<td>127</td>
<td>2/0.02</td>
</tr>
<tr>
<td>TOPICS</td>
<td>125</td>
<td>6/0.05</td>
</tr>
<tr>
<td>Grand Total</td>
<td>925</td>
<td>52/0.06</td>
</tr>
</tbody>
</table>

and parameters for the coding are presented in Table 2.

Results and Discussion of Literature Review

It is noteworthy that, out of 925 articles published in a sample of seven journals in special education across a recent 5-year span, only .06% involved international collaboration. The topics of collaboration were varied and, due to the single subject parameters of the review, focused on behavioral interventions that had an effect on the academic or functional skills, social behaviors, or communication skills of students with disabilities across types of disability and age level. Faculty members participating in the collaborative research were from a small number of countries, and it is not possible to determine their roles in the research (e.g., conceptualizing research, implementing study, analyzing results, translating manuscript for submission). It also is not clear how these collaborations came into being, but it is possible that doctoral study experience or similar research agendas could have played a role. This limited review of collaborative special education research across countries is only a snapshot and is not intended to reflect the percentage of collaborative research published across all special education journals published in English. It does, however, provide evidence of a paucity of international collaboration in the field of special education and sets the stage for the following description of the impact of a collaborative relationship between special education researchers from Turkey and the United States.

Example of an International Collaboration

In the following sections, we provide the personal narratives of two faculty members (first and second authors) who have been involved in international collaboration. The purpose of this section is to demonstrate the impact that international collaboration can have on the research, development, and dissemination of evidence-based practices in special education.
<table>
<thead>
<tr>
<th>Journals</th>
<th>Authors</th>
<th>Topics</th>
<th>Countries</th>
<th>SS Research Method</th>
<th>Funding Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETADD (2015)</td>
<td>Seok, DaCosta, &amp; Yu</td>
<td>Effects of spelling practice intervention using tablet computer and picture cards</td>
<td>Korea, USA</td>
<td>ATD with a non-concurrent MBD</td>
<td>National Research Foundation in Korea</td>
</tr>
<tr>
<td>ETADD (2013)</td>
<td>Goh &amp; Bambara</td>
<td>Effects of video-self modeling in teaching chained tasks</td>
<td>Singapore, USA</td>
<td>MPD across behaviors</td>
<td>N/A</td>
</tr>
<tr>
<td>ETADD (2011)</td>
<td>Rispoli et al.</td>
<td>Effects of pre-session satiation on challenging behavior and academic engagement</td>
<td>New Zealand, USA-Italy</td>
<td>ATD</td>
<td>N/A</td>
</tr>
<tr>
<td>ETADD (2011)</td>
<td>Zisimopoulos, Sigafoos, &amp; Koutromanos</td>
<td>Effects of video prompting and time delay procedures to teach basic internet search</td>
<td>Greece-Australia</td>
<td>MBD across participants</td>
<td>N/A</td>
</tr>
<tr>
<td>ETADD (2010)</td>
<td>Ohtake, Wehmeyer, Uchida, Nakaya, &amp; Yanagihara</td>
<td>Effects of a time delay procedure on the acquisition of skills for repairing multiple listeners' misunderstandings</td>
<td>Japan, USA</td>
<td>CCD</td>
<td>N/A</td>
</tr>
<tr>
<td>ETADD (2010)</td>
<td>Yilmaz, Konukman, Birkan, &amp; Yanardag</td>
<td>Effects of a most to least prompting procedure on teaching swimming skills</td>
<td>Turkey, USA</td>
<td>MPD across participants</td>
<td>N/A</td>
</tr>
<tr>
<td>ETADD (2010)</td>
<td>Yilmaz et al.</td>
<td>Effects of a constant time delay procedure on swimming rotation skills</td>
<td>Turkey, USA</td>
<td>MPD across behaviors</td>
<td>N/A</td>
</tr>
<tr>
<td>TESE (2013)</td>
<td>Barton, Chen, Pribble, Pomes, &amp; Kim</td>
<td>Effects of training/coaching on preservice teachers in teaching play skills</td>
<td>USA-Chile, Korea</td>
<td>MBD across participants</td>
<td>N/A</td>
</tr>
<tr>
<td>TOPICS (2011)</td>
<td>Blair, Lee, &amp; Dunlap</td>
<td>Impact of behavior support on the children’s behaviors</td>
<td>Korea, USA</td>
<td>Concurrent MBD across participants</td>
<td>N/A</td>
</tr>
</tbody>
</table>

ATD: Alternating treatments design, MPD: Multiple probe design, MBD: Multiple baseline design, CCD: Changing criterion design.
As a doctoral student at the University of Kentucky (UK), I was trained as a researcher in the area of systematic instruction using response-prompting strategies and worked on several federally funded research grants under Drs. David Gast and Mark Wolery that used single subject research methodology to examine the strategies in small group instruction, to compare strategies, and to assess their effectiveness in teaching safety skills. While this research built on the initial research of others in the field, response-prompting strategies became associated with UK through publications (e.g., Wolery, Ault, & Doyle, 1992) and presentations. When I later became a faculty member at UK, I continued this line of research with colleagues and graduate students. Our continuing line of research questions led to establishing response-prompting procedures (e.g., system of least prompts, constant time delay, simultaneous prompting) as evidence-based practices in teaching students with disabilities while also building evidence for the addition of nontargeted information, the use of observational learning, and a number of generalization strategies as best practices to increase the efficiency of instruction. Across 25 years, I was privileged to be involved in 54 studies on systematic instruction published across 12 journals, as well as book chapters and books (e.g., Collins, 2012).

As faculty and doctoral students at UK left for faculty positions at other Institutions of Higher Education (IHEs), the line of research on systematic instruction resulted in further investigations at other IHEs that contributed to the professional literature with innovative research questions. In addition, colleagues at other IHEs across the country also designed, completed, and published independent investigations in the area of systematic instruction, adding the names of their colleagues and students to the research base.

While it is not unusual for special education practices that are disseminated through professional publications and presentations to be cited and used to ask new questions across the United States, the practice of international collaboration has the potential to have a global impact on the field special education. This global impact is demonstrated through an international collaboration with a faculty member (Elif Tekin-Iftar) from Turkey that I met while I was a faculty member at UK.

In 1996, as a doctoral student, I attended the annual conference of the Council for Exceptional Children (CEC) in Orlando, FL, with my doctoral advisor, Dr. Gonul Kircaali-Iftar. I had decided to base my research career on single-subject research methods because this methodology allows the analysis of individual differences while establishing strong experimental control between dependent and independent variables; as a result, researchers may provide more information about the effectiveness of procedures with those who have learning or behavioral challenges. During the conference, I attended a session where Dr. Edward Blackhurst from UK introduced the “Single Subject Research Advisory Program” developed at UK to provide computer-based decision-making for selecting appropriate single subject research designs to answer applied research questions in the field of special education. Special education was an emerging field at that time in Turkey, so I decided to go abroad to study current behavioral and instructional strategies and identified UK as a place that matched my research interests. I received a scholarship from the Turkish Academy of Sciences and pursued part of my doctoral education at UK between 1997–1998 (17 months). While I was there, I prepared my dissertation proposal, audited courses (e.g., Applied Behavior Analysis, Systematic Instruction for Students with Moderate to Severe Disabilities), and took part in several research activities with Dr. John W. Schuster as my mentor during my visit. In addition, I studied with several faculty who also had an interest in severe disabilities, ABA, single subject research design, and systematic instruction, including the co-author of this paper, Dr. Belva C. Collins.

When I returned to Turkey, I implemented research on the efficacy of response-prompting procedures delivered by sibling tutors to teach discrete skills to their siblings with intel-
lectual disabilities. I found it difficult to find adequate resources on systematic instruction in the Turkish research literature and relied on much of the literature coming from UK on this topic. When I finished my doctoral degree in 1999, I began to conduct research projects focusing on response-prompting strategies with my students and colleagues and to teach graduate level courses on systematic instruction. Almost one year later, I returned to UK to conduct a literature review for a book on errorless teaching since open resources and access to journals were limited in Turkey at that time. In 2001, my doctoral advisor and I published a book entitled *Errorless Teaching Methods in Special Education* (Ozel Egitimde Yansiz Ogrenim Yontemleri), and an article based on my dissertation was published in *Education and Training in Mental Retardation and Developmental Disabilities* (Tekin & Kircaali-Iftar, 2002).

In 1999, I became a faculty member at Anadolu University. During the subsequent years, my graduate students conducted their thesis and dissertation research on response-prompting strategies, and we presented these studies at international and national conferences (i.e., Association for Applied Behavior Analysis International, National Special Education Conference in Turkey). We also published our research in various international and national journals, such as *Autism, Education and Training in Autism and Developmental Disabilities, Research in Developmental Disabilities,* and *Topics in Early Childhood Special Education* (Akmanoglu & Tekin-Iftar, 2011; Dogan & Tekin-Iftar, 2002; Ersoy, Tekin-Iftar, & Kircaali-Iftar, 2009; Kurt & Tekin-Iftar, 2008; Yildirim & Tekin-Iftar, 2002). At the same time, I organized several research teams, conducted independent research projects, and published the results (nine articles) in several international journals (e.g., *Development and Education, International Journal of Disability, The Journal of Special Education*). In addition to research with my students, colleagues in Turkey who are either at the same university (e.g., Batu, 2008; 2014; Birkan, 2005; Celik & Vuran, 2014; Ozen, 2008; Vuran, 2008) or other universities (e.g., Aykut, 2012; Ciftci-Dere & Temel, 2010; Karabulut & Yikmis, 2010) in Turkey also started to conduct independent research projects and/or supervise thesis and dissertation research on systematic instruction using response-prompting procedures (e.g., Ari, 2008; Kanpolat, 2008; Karisyakali, 2011; Onal, 2008; Ozbey, 2005; Yeterge, 2015). My colleagues and I found that faculty from different disciplines (i.e., physical education, physiotherapy, music education, computer education and instructional technologies) that also serve children with disabilities noticed our efforts in this line of research, and they started to conduct research projects on the same topic. Thus, international and national publications in related disciplines have emerged (e.g., Eren, 2012; Gokmen, Tekinarslan, & Ciftci-Tekinarslan, 2015; Yanardag, Birkan, Yilmaz, Lieberman, & Agbuga, 2011; Yanardag, Mert, Yilmaz, Arican, & Duzkantar, 2015; Yilmaz, Birkan, Konukman, & Erkan, 2005; Yilmaz et al., 2010). As a result of these cumulative efforts, the number of research articles published by faculty in Turkey on systematic instruction for individuals with disabilities has increased since 1999, adding to the international research base on the topic and extending the identification of evidence-based practices across geographic regions and researchers.

The increase in the number of researchers studying these topics and publishing in this line of research has resulted in citations across researchers in Turkey and the US, creating a scaffolding effect where studies from the two countries have built on each other’s work as researchers from both countries pursued similar lines of inquiry. The effects of the initial international collaboration also have been observed in translating research into practice since almost all of the special education departments (undergraduate and graduate level) in different universities in Turkey (and in Cyprus) have started to include response-prompting strategies in their systematic instruction course content. Therefore, many special education teacher candidates in Turkey have acquired competency in using these strategies while working with their students with developmental disabilities, as is the case in the US. Figure 1 displays the special education departments in Turkey where response prompting strategies are included in their special education teacher training curricula.

Continuing collaboration with colleagues at UK has given us the opportunity to organize
several professional meetings in Turkey. One UK professor (Dr. Michael Nelson) came to the first International Conference in Special Education in Turkey where he made a speech and delivered workshops about positive behavior support in 2001. Two others that I met while at UK (Drs. Jennifer Grisham-Brown and Mary Louise Hemmeter) conducted a series of workshops and conferences about naturalistic teaching approaches for children with developmental disabilities and models of professional development for training special education teachers in 2012. All these efforts have provided a base for teachers to implement evidence-based practices in their classrooms and graduate students to update themselves on the current literature on these topics.

Almost two decades following my first collaboration with UK, I have established a second collaboration in the US. I received a post-doctoral fellowship from The Scientific and Technological Research Council of Turkey (TUBITAK) to come to the University of North Carolina at Charlotte (UNCC) as a visiting scholar, where Dr. Belva Collins (former professor and department chair at UK) is a current professor and department chair. At the same time, three of my doctoral students have received funding from TUBITAK and their university to be visiting scholars: one at UNCC and two at UK and Vanderbilt. Thus, as a group of researchers from Turkey, we have started a second-generation collaboration with researchers in the US, stemming from my initial collaboration with UK. Figure 2 displays the timeline of my collaboration with faculty in the US.

This collaborative journey is shared to explain how it is possible for an international collaboration to contribute to and support evidence-based practices in the field of special education. Over the past 20 years, my colleagues and I have trained our graduate students in Turkey to use the practices I learned in my initial collaboration to pursue this line of research. Some are now our colleagues and have started to train their students to pursue the same line of research. Colleagues currently supervise thesis and dissertation research (e.g., Genc, 2010) and conduct independent research projects using single subject research methodology to further investigate the parameters of systematic instruction with students with disabilities (Akmanoglu, Kurt, & Kapan, 2015; Odluuyurt, 2011; Ulke-Kurkuoglu, 2015; Yucsesoy-Ozkan & Gursel, 2011). Table 3 displays the number of theses and dissertations conducted in Turkey that have examined the effects of response-prompting strategies. These numbers are more meaningful considering that only six universities in Turkey provide master’s degree programs (see Figure 1), and only three of these also provide doctoral programs in special education; thus, the number of graduate students in

Figure 1. The cities where research about prompting strategies and graduate studies are conducted (open asteriks) and where prompting strategies are taught in teacher training programs (black asteriks). Note: Two universities offer graduate studies in special education in Ankara.
Turkey is limited, with each program accepting between 5–10 students per year. In light of this, it can be argued that response-prompting strategies are one of the most frequently studied topics across different geographical areas in Turkey.

Another indicator of the contribution of our international collaboration is the number of publications on response-prompting strategies by faculty and students from Turkey. As seen in Tables 3 and 4, the most frequently investigated topics have been variations of the simultaneous prompting and constant time delay procedures, also across different geographical areas in Turkey.

The same teaching and research efforts have continued with our colleagues in the US. It is common for professors across Turkey and the US to require their students to read research studies on systematic instruction from both countries, to cite these studies in their own lines of continuing research, and to work with students to develop and implement single subject research studies on response-prompting strategies that build on the published research from both countries. The impact that this has on teachers in the field across Turkey and the US cannot be measured, but the requirement to use evidence-based practices in classroom instruction guarantees that the application of research to practice is occurring.

### Discussion of International Collaboration

As demonstrated in the opening section of this article, international collaboration in special education is evident in some of the research studies published in the sample of journals that we reviewed. Each time a study is

---

**TABLE 3**

<table>
<thead>
<tr>
<th>Prompting Strategies</th>
<th>Theses</th>
<th>Dissertations</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous prompting</td>
<td>25</td>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td>Constant time delay</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Most to least prompting</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Least to most prompting</td>
<td>1</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Graduated guidance</td>
<td>–</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Antecedent prompt and testing</td>
<td>3</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>Grand total</td>
<td>42</td>
<td>10</td>
<td>52</td>
</tr>
</tbody>
</table>
published in an internationally distributed peer-reviewed journal, the opportunity is present for practices to become evidence-based (i.e., be replicated across geographic locations by different researchers) and for research-to-practice to increase globally.

The sharing of our personal stories illustrates the impact that a single initial collaboration can have on the development of a single line of research and the subsequent impact that this can have on global special education practices. We, therefore, offer the following considerations for those who are considering establishing an international collaboration in research and practice.

**International Perspective from the United States**

The establishment of an international research collaboration can come about in several ways. If a doctoral program accepts international students, the research collaboration can begin during the student’s tenure in the program. Even if a doctoral student does not apply for formal admission to a doctoral program in the US, it is possible to bring an international student to a university as a visiting international scholar to study a specific topic with researchers without enrollment in coursework. In the case of Dr. Tekin-Iftar, UK extended an invitation to her to work with faculty in the department as a visiting scholar. While at UK, she sat in on classes taught by the sponsoring researchers and observed faculty as they performed departmental responsibilities (e.g., clinical supervision) and conducted research while receiving feedback on her own work and developing her own research agenda. One way to begin a collaboration is to network at research conferences attended by an international audience. This is how Dr. Tekin-Iftar made her first contact with faculty from UK. Sitting in on the presentations of international researchers can lead to further dialogue on research collaboration. Even if international researchers cannot meet face to face, they can follow each other’s published work and establish a dialogue using direct technology (e.g., email, videoconferencing, web conferencing) or through an established virtual network (e.g., Researchgate).

In searching for a collaborative relationship with an IHE, scholars need to consider both the clinical and research needs of their countries. While experienced researchers should have this information and perspective, it is vital for younger researchers to pay attention to this. The roles and perspectives of the experienced advisors are important as they use their experiences to mentor a younger generation of beginning faculty. This point is validated in the case of Dr. Tekin-Iftar. The perspectives and support of her advisor (Dr. Gonul Kircali-Iftar) enabled her to begin her initial collaboration, and she continues to mentor her own doctoral students in international collaboration. Therefore, we strongly suggest that younger researchers and graduate students discuss the option of international collaborative experiences with the experienced researchers who mentor them and get their thoughts and ideas on the benefits of

### Table 4

<table>
<thead>
<tr>
<th>Prompting Strategies</th>
<th>Number of Articles</th>
<th>Turkish City Where Study was Implemented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous prompting</td>
<td>16</td>
<td>Eskisehir ( (n = 14) )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ankara ( (n = 2) )</td>
</tr>
<tr>
<td>Constant time delay</td>
<td>10</td>
<td>Eskisehir ( (n = 9) )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ankara ( (n = 1) )</td>
</tr>
<tr>
<td>Most to least prompting</td>
<td>4</td>
<td>Eskisehir ( (n = 3) ),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Istanbul ( (n = 1) )</td>
</tr>
<tr>
<td>Antecedent prompt and test</td>
<td>2</td>
<td>Eskisehir ( (n = 2) )</td>
</tr>
<tr>
<td>Graduated guidance</td>
<td>2</td>
<td>Eskisehir ( (n = 2) )</td>
</tr>
<tr>
<td>Least to most prompting</td>
<td>2</td>
<td>Eskisehir ( (n = 2) )</td>
</tr>
</tbody>
</table>
this type of experience. Their perspective might open a door to a productive future career and enable them to use financial resources available to them in a responsible way.

Department chairs often receive email requests from international students and faculty who want to come study (formally or informally) at an IHE in the US. Those who are successful most often spend time researching the vitae of faculty in a department to see if there is a match in research interests and make an effort to read articles and attend conferences where the researcher’s work is disseminated. In addition, a well-written letter stating research goals and a curriculum vita that shows a research agenda help in establishing a possible collaboration. Still, the collaboration is most likely to develop through personal contact or through a personal network (e.g., personally known international scholar recommends a student or faculty member to a department).

Once the contact is made and the door of opportunity is opened, researchers must decide how to proceed. Informal short-term visits can be set up to meet and discuss possible research projects. For example, a student or faculty member from either country may locate short-term travel funding (e.g., university stipend, scholarship from research organization) to travel to a country for a mutually agreed time period to discuss and design research. In this case, it is up to the person traveling to make all arrangements and assume responsibility for the costs of visiting another country (e.g., passport, airline tickets, rental car, daily per diem for meals) although the host may be able to assist with accommodations.

If the collaboration is to be more formal in nature and occur over an extended time, there will be responsibilities for both parties. For example, bringing a visiting scholar to an IHE in the US may involve having the host complete paperwork for the IHE (e.g., letter of invitation, documentation that the visitor speaks English), locate office space, provide a computer with email and library access, secure signatures on documents to obtain a visa, and complete modules on international research through the Institutional Review Board. It is important to consult the IHE’s international office as soon as a visit is discussed since the host’s responsibilities may take some time to complete. The visitor also has responsibilities that include securing a passport and visa and locating funding for the trip. While requirements for a face-to-face collaboration can be checked off a list, there are additional considerations that may include obtaining necessary documentation for anyone who will accompany the scholar, making plans for funds in the new country (e.g., credit card, cash, bank account), identifying a mode of transportation (e.g., public transportation, rental or purchased car using international driver’s license), securing housing for the duration of the visit (e.g., extended stay hotel, furnished rental apartment, room in a home), and ensuring that medical needs are covered (e.g., health insurance policy). Finally, it is always wise to know the locations of the native country’s embassy and to register with the international travel branch of the home country in the event of an international crisis that may threaten the safety of a traveler; information about safeguards can be accessed while securing a passport or visa.

It also is possible that collaborating researchers will want to bring students or colleagues with them. Formal options through the university should be investigated since going through these channels puts students under the protection of the university’s liability insurance. Most students enroll in course credit under a traveling faculty member, which involves the payment of tuition in addition to other costs associated with the trip, including health/travel insurance. The researcher will want to provide a detailed agenda of the trip, including accommodations and a travel itinerary. If the travel is funded through the IHE, it may be possible to obtain a time-limited credit card through the university to pay for accommodations, meals, tours, and other associated expenses.

Last, both collaborating researchers need to pay attention to the sustainability of the collaboration they have created. Information obtained through research changes and evolves through the years; therefore, any collaboration should not be limited to one interaction. Instead, sustainability of the collaboration over the years involving a new generation of younger researchers should be planned. In our example, we are optimistic that our group
of younger researchers will provide impact data on evidence-based practices in our field and continue the relationships they have established.

Finally, it is important to note that speaking and writing in a second language for prestigious journals is always a barrier for non-English speaking researchers. Visiting scholars should consider developing their speaking and writing skills (not only for publication) prior to arriving in a new country. In respect to writing in a second language for prestigious journals, journal editors and reviewers need to consider the research quality of the manuscript and try to provide feedback about the use of language and grammar in the manuscript in a positive way. Journal editors may consider providing editing services to the greatest extent possible.

When reviewing the literature to find the research articles written through international collaboration, we only located a few articles in our journal sample. The possible reason for not having more articles can be due to the lack of the researchers’ international experiences, finding sponsorship to make the collaboration possible, not having an international perspective, and the nature of educational/social science fields (e.g., national policies and legal regulations that differ across countries).

Conclusion

In spite of the upfront planning that must be completed prior to traveling abroad to collaborate on research, the outcomes well outweigh the cost and the work. International collaboration allows researchers to develop a line of research that continues for decades. It allows them to form a relationship that can involve current and future students. It allows the field to benefit from multiple lines of inquiry that validate and refine research results. In the field of special education especially, it allows global services to be improved and the lives of children with disabilities to be changed through the establishment and implementation of evidence-based practices.

References


Collins, B. C. (2012). *Systematic instruction for students*


Karabulut, A., & Yikmis, A. (2010). The effectiveness of simultaneous prompting on teaching the skill of telling the time to individuals with mental retardation. Abant Izzet Bayhas Universities Eğitim Fakultesi Dergisi (Journal of Abant Izzet Bayhas University Faculty of Education), 10, 103–113.


Onal, Z. K. (2008). Akran öğrencilerle destek ile sunulan sabit bekleme süreli öğretimin genel eğitim sınıflarında etkiliği (The investigation of the effectiveness of peer mediated...


Received: 28 July 2016
Initial Acceptance: 27 September 2016
Final Acceptance: 15 November 2016
Early Intervention in Autism Spectrum Disorder: Concerns and Support of Portuguese Mothers

Lúcia S. Magalhães and Ana P. S. Pereira
University of Minho

Abstract: Families of children with autism spectrum disorder (ASD) face a series of daily challenges, from which emerge a number of concerns that are inherent to the reality of living with a child with ASD, as well as the demand for formal and informal supports that are required for these families. This study’s main objectives were to: (1) identify and consider the concerns of Portuguese mothers of children with ASD; (2) identify and consider the required forms of support that these mothers need; and (3) understand if the support networks that are given to these mothers are adequate in face of their concerns. Results showed that these mothers hold concerns and support that are consistent with most of the scientific studies in this area. Results also indicate a greater demand for formal support throughout the life of the child with ASD, rather than informal support. The inadequacy of support provided by the primary health care system is highlighted by the mothers, especially at the initial stage and in the diagnostic process. Finally, results suggest the adequacy of specialized formal support, especially in the context of early intervention, although mothers complain about the amount of existing supports, as well as the frequency and duration of sessions provided by therapeutic services and early intervention programmes.

Early Childhood Intervention in Portugal

Early Childhood Intervention National System (SNIPI)

SNIPI is an organized group of institutional entities that act to ensure development conditions in education, and in the health and social system, for children aged 0 to 6 years who have changes in body functions or body structures, with severe risk of developmental delays, and their families (Decree-law No. 281/09, of October 6, 2009).

With regard to eligibility criteria, according to Decree-law No. 281/09, of October 6 (2009), children between 0 and 6 years and their families are eligible for support under the SNIPI if they present the following conditions: (1) changes in body functions or structures; or (2) high risk of developmental delay. All children in group 1 and group 2 who have four or more factors of biological and/or environmental risk are eligible for access to SNIPI.

SNIPI works by joint representative structures of the Ministries of Labour and Social Solidarity, Health and Education, in close collaboration with families, and is coordinated by the Coordination Committee of SNIPI. The aims of SNIPI are to: a) ensure the protection of children’s rights and develop their skills through Early Intervention (EI) actions throughout the country; b) detect and flag all children at risk of changes or with effective changes in body functions and structures, as well as children with serious risk of developmental delay; c) intervene according to the needs of the family, in order to prevent or reduce the risk of developmental delay of the child; d) help families in accessing services and resources related to health, education and social systems; and e) involve the community by creating articulated mechanisms of social support (Decree-law No. 281/09, of October 6, 2009).

EI practices aim to promote and ensure the participation and social inclusion, as well as
the quality of life of these children and their families, for which access and response must be universal. To ensure and implement the principles adopted by the SNIPI, there are EI local teams.

EI local teams operate at the level of municipalities or parishes and are composed of doctors, nurses, psychologists, therapists, social workers, and kindergarten teachers who work for the Ministries of Labour and Social Solidarity, Health and Education. They are responsible for: a) identifying children and families who are eligible for SNIPI; b) ensuring the care needed by children and families who, although not immediately eligible for SNIPI, require periodic evaluation; c) helping ineligible children and families who need social support; d) preparing and performing the individual EI plan for each child; e) identifying the needs and resources of the community, mobilizing formal and informal forms of support; f) working in collaboration, when appropriate, with the committee for the protection of children and young people and/or with other entities who work in the field of child protection; g) ensuring, when necessary, suitable transitions from the SNIPI to other programs, services or educational settings; and h) collaborating with kindergarten and nursery teachers (Decree-law No. 281/09, of October 6, 2009).

Families of Children with Autism Spectrum Disorder

Research in the area of ASD has sought to understand and explore its influence on families and on each family member, their functions and dynamics, as well as in their interactions with employment, leisure, the extended family, and the community. The majority of studies has demonstrated a relationship between ASD and the health of parents, both at the physical and psychological level, where it is possible to find high levels of depression, feelings of disbelief, pain, anxiety, guilt and anger, as well as low levels of physical and mental well-being (Baker-Ericzén, Brookman-Frazee, & Stahmer, 2005; Hoffman, Sweeney, Hodge, Lopez-Wagner, & Looney, 2009; Karst & Van Hecke, 2012).

Literature has shown that these families experience high levels of stress (Harper, Dyches, Harper, Roper, & South, 2013; Hoffman et al., 2009) which can be triggered by diverse factors. The most common triggers are a variety of concerns related to the experience and care of the child with ASD, namely:

- The child’s inappropriate and unpredictable behaviour (Hall & Graff, 2011),
- Problems regarding social and parental relationships with the child (Hoffman et al., 2009),
- The diagnostic process (Stuart & McGrew, 2009),
- Lack of knowledge and information regarding ASD (Stuart & McGrew, 2009),
- Overload of care regarding the child’s needs (Harper et al., 2013),
- Concerns about the child’s future and education (Dillenburger, Keenan, Doherty, Byrne, & Gallagher, 2010),
- Lack of and demand for various resources and social support (Dillenburger et al., 2010; Harper et al., 2013),
- Poor acceptance of the child with ASD by the extended family and community (Sullivan, Winograd, Verkuilen, & Fish, 2012), and
- Unmet needs (Dillenburger et al., 2010).

The characteristics associated with ASD cause major limitations in the functioning and the dynamics of the family system, as well as the interactions between children with ASD and their families. The lack of reciprocity and emotional involvement of the child, combined with family difficulties in responding to and comforting the child, greatly influence the quality of the interaction between parents and their children, resulting in a lower sense of parenting and parental competence (Hoffman et al., 2009; Karst & Van Hecke, 2012).

Simultaneously, changes in marital and family life, poor adaptability to the problem of raising a child with ASD, difficulties in family cohesion, and expressions of affection are common in these families (Harper et al., 2013; Nealy, O’Hare, Powers, & Swick, 2012).

Also, the relationship between parents and the siblings of children with ASD may be further compromised and neglected by the immense time and care that the child with ASD requires. Regarding these siblings, research shows that they feel neglected, alone and con-
cerned about the future of the child with ASD, in comparison with siblings of typically developing children (Green, 2013).

Adherence to rigid and inflexible routines by the child, combined with the difficulty in accepting change, create numerous alterations in the family’s routines and in the family’s involvement in the community, implying several adaptations with respect to all members of the family (Lee, Harrington, Louie, & Newschaffer, 2008; Nealy et al., 2012). As a result, these families tend to isolate themselves from the contexts in which they live, choosing not to participate in activities with their extended family, recreational and leisure groups, and the rest of the community (Dillenburger et al., 2010). In such contexts, parents also referred to the great need to provide information and educate the community about ASD, in order to protect their children from prejudice and hostile behavior (Sullivan et al., 2012).

The care that children with ASD require implies a great expenditure of time and energy by families, causing alterations in family routines, including changes in family management and schedules, which also compromises future possibilities of job promotion (Lee et al., 2008; Nealy et al., 2012).

Most studies also report that there are grave concerns about the future of the child with ASD, depending on the child’s development stage and the challenges that are posed to both child and family. Starting school and the transition from pre-school to primary education are a source of great concern for most parents (Dillenburger et al., 2010; Karst & Van Hecke, 2012).

Recent research has sought to understand and explore this topic among parents of children with ASD, making it possible to identify and categorize the main concerns felt by families, namely: (a) concerns associated with the diagnostic process (Siklos & Kerns, 2007); (b) economic concerns (Hall, 2012; Nealy et al., 2012; Siklos & Kerns, 2006); (c) concerns regarding the type, quality and functioning of available services and resources (Siklos & Kerns, 2007); (d) information requirements (Renty & Roeyers, 2006); (e) recreation and socialization needs (Nealy et al., 2012); and (e) unmet needs (Siklos & Kerns, 2006).

Current Study

Moreover, research has shown the importance of EI in the response to the concerns of families of children with ASD (Reis, Pereira, & Almeida, 2014). The demand and implementation of support as early as possible for the child with ASD implies an increase in the concerns of the family system. Therefore, this step is essential to carry out intervention programs focused on child development and on responding to the concerns of families, by mobilizing support and helping the family adapt to the challenges and demands associated with caring for a child with ASD (Reis et al., 2014). However, research also shows that professionals have a tendency to focus intervention only on the child with ASD, while not appreciating the concerns and priorities of families (Bultas, 2012). Just a few Portuguese studies examined specifically the concerns and required supports of families of children with ASD. In the absence of enough research, this field has little understanding of the functioning and adequacy of formal and informal supports that are required for these families.

Therefore, the aims of this study were to: (1) identify and recognize the concerns of Portuguese mothers of children with ASD; (2) identify and recognize the required means of support those mothers need; and (3) understand the adequacy of the support networks given to these mothers, according to their concerns. As a result, three questions guided our work:

1. Which are the concerns of Portuguese mothers of children with ASD?
2. Which resources and supports do these mothers have access to?
3. How do these resources and supports operate, in face to these mothers’ concerns?

To answer these questions, we followed a qualitative methodology study with 12 Portuguese mothers of children with ASD, ages 3 to 6 years old. The qualitative methodology adopted is in line with the questions guiding this investigation, considering the need for a more detailed, in-depth, individual and personal research, on a subject which research is emerging (Bogdan & Biklen, 1994).
Method

Participants

Participants were Portuguese mothers of children with ASD, ages 3 to 6 years old, enrolled in the SNIP. These mothers are the family members who spend more time with the child, accompanying him/her in seeking and obtaining support (e.g., medical and therapeutic support; schooling; access, transport and monitoring services and resources; basic care). Exclusion criteria included mothers of children with other developmental disorders besides ASD and/or children without a formal diagnosis of ASD.

After requesting the collaboration of an institution specialized in supporting families and children with ASD, 20 mothers who fulfilled the inclusion criteria were individually asked to participate in the study. The 12 mothers who consented to participate signed an informed consent. All of these mothers lived in the same urban district and were supported by an institution that is specialized in supporting families and children with ASD, with two psychologists, two speech therapists, two occupational therapists, one child psychiatrist and one social worker. This team gave EI support, although it did not belong to the SNIP.

Following, participants were contacted by telephone by the investigator to schedule the interviews at a time chosen by each mother. All the interviews were carried out in the institution where the families and their children were supported, in a room familiar to the participants, in order to ensure a safe and non-threatening natural context for the mothers and to create an environment favorable to share information (Bogdan & Biklen, 1994).

The selection of mothers was performed according to an intentional selection criterion, by integrating subjects we considered to be an asset to the collection of as much relevant information as possible to the subject and study objectives (Patton, 2002). The mothers were selected in series, taking into account the data obtained with the above mentioned subject, and looking for different information and/or clarifying certain aspects with the following subjects.

Data Collection Instrument

The instrument used to collect data was a single semi-structured interview. Previously, the preparation of a preliminary interview guide was made, followed by its validation. This validation included the interview of a mother followed by an in-depth analysis by a specialist in qualitative research. This process provided a review of the content and form of the questions, the language that was used, as well as the effectiveness of the interview script, in light of the purposes of the research. Finally, we achieved a final guide with a set of open-ended questions in a sequence that was appropriate to the study objectives (Kvale, 1996).

Each interview began with the presentation of relevant information concerning the subject and the objectives of the study. The first three questions were focused on bringing the mothers to the main themes, and also, on fostering an empathic and friendly atmosphere to their participation (Kvale, 1996). Then, specific direct questions were made to search for data related to the study phenomenon, in a flexible way, according to the course of the interview. Whenever necessary, follow-up questions, probing questions, indirect, structuring and interpreting questions were performed (Kvale, 1996). Interviews were completed with a brief summary of the main information obtained, giving opportunity for mothers to add information and/or request clarification concerning any possible doubts (Kvale, 1996). All the interviews were performed by one same investigator, and each took about 45 minutes. Data was collected using a digital recorder, and later on transcribed and computerized by the investigator.

Data Analysis

All data from interviews was converted into a single data file. First, these transcripts were read by the investigator to create the material for analysis, and emergent categories were created using deductive and inductive methods (Bardin, 2009; Lima, 2013). To ensure the material’s reliability and validity, two coders, both familiar and holding theoretical and practical knowledge about ASD, coded the
TABLE 1
Category System

<table>
<thead>
<tr>
<th>1. Concerns of the mothers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Support for the mothers</td>
</tr>
<tr>
<td>2.1. Formal support</td>
</tr>
<tr>
<td>2.2. Informal support</td>
</tr>
<tr>
<td>3. Adequacy of the forms of support in light of the concerns of the mothers</td>
</tr>
<tr>
<td>3.1. Formal support response</td>
</tr>
<tr>
<td>3.2. Informal support response</td>
</tr>
</tbody>
</table>

data independently of the categories defined by the investigator.

The percentage agreement index (i.e., number of agreements divided by number of agreements plus disagreements and multiplied by 100) was used to calculate intercoder agreement. According to some authors, agreement percentages above 80% are considered good indicators of reliability, although others point more demanding values, such as 85% (Coutinho, 2011). The calculated percentage showed a substantial rate of agreement, as we obtained an intercoder reliability above 80 (89.79%) (Krippendorf, 1980; Lima, 2013). Therefore, a final coding was performed, organizing the units in different categories and subcategories.

Results

The presentation of the results was carried out taking into account the research’s empirical questions, and the category systems taken from the content analysis of the research data (Bardin, 2009; Bogdan & Biklen, 1994).

Concerns of the Mothers

Mothers mentioned concerns at different times of their lives and, especially, the lives of their children. These concerns were related to their own child, the family system, and broader aspects such as the social and economic environment.

The most common difficulties mentioned by the mothers were the child’s lack of speech and eye contact. Participants reported that these concerns appeared when their children were 18 months. Three mothers referred the presence of restrictive and repetitive behavior, interests and activities. For example, one mother said, “he was constantly fascinated with the washing machine”.

Most mothers highlighted the presence of feelings of anxiety, shock, guilt and non-acceptance of the diagnosis of ASD, especially after its confirmation. One mother described her feelings in this way: “it’s so difficult and hard . . . later, you blame yourself because you think it’s your fault, you didn’t do things correctly or, in the past you did something and now, you are being punished. So you ask yourself why? Why was I chosen to have a child with this problem?”. However, a minority of participants shared a different view regarding the diagnosis. For example, a mother said “we don’t see ourselves in another situation. For us, it’s already normal ( . . . ) More work? Yes! But at this moment, we don’t see ourselves in another situation”. Another mother said that the diagnosis gave her a new vision about her own life and the resolution of some problems: “Now it’s like this: I now have my own things to think about. Things we used worry about before, are no longer important”.

Some mothers said they had concerns and expectations regarding their child’s development and future, especially about his/her autonomy and future care in the event of the death of a parent. For example, one mother said, “the future worries me . . . what’s he going to do when I die? Who will take care of him?”

The majority of mothers said they had needed information about ASD when they first received the diagnosis, as one mother said, “when I was told that my child had autism, I asked, “what’s that?”. I didn’t know nothing about that”. Mothers also mentioned the need for information regarding the necessary support available after the initial diagnosis: “At first, I asked some questions and doctors were not very clear. I had many doubts . . .” and “another thing I didn’t like was that when I was given the diagnosis, I was left on my own”.

Half of the mothers said they had many concerns related to the daily functioning of the child, as one mother explained: “There is always a battle: the pacifier, the tantrums, mood swings, now it’s a battle with potty training . . . there are always challenges . . .”.

TABLE 1
Category System

<table>
<thead>
<tr>
<th>1. Concerns of the mothers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Support for the mothers</td>
</tr>
<tr>
<td>2.1. Formal support</td>
</tr>
<tr>
<td>2.2. Informal support</td>
</tr>
<tr>
<td>3. Adequacy of the forms of support in light of the concerns of the mothers</td>
</tr>
<tr>
<td>3.1. Formal support response</td>
</tr>
<tr>
<td>3.2. Informal support response</td>
</tr>
</tbody>
</table>
The majority of mothers shared their concern and anxiety regarding the transition to the first year of school: “He still has one year and a half before starting primary school. If only he could communicate and integrate more even if he doesn’t speak...”. Another mother said: “My fear is when he goes to school. Will he be treated badly? Will they hurt my child?”

The burden of care was a concern shared by half of the participants. One mother stressed: “And then I do nothing! I am always with my son running from one place to the other. That’s my life”. Some mothers spoke about the difficulties in reconciling their professional life with the role of parenting a child with ASD. Here, some mothers said: “I can’t come more often to therapeutic services. Our jobs don’t allow us that, unfortunately”, “I think that mothers do not have the chance to hold a full-time job and, at times, we have to hide this fact in order to get a job”.

Mothers reported the need for psychological support, especially after the diagnosis. For example, one mother said: “I think there should be more support like psychological help for parents because having a child with ASD is a big blow”. These mothers highlighted the importance of psychological help, not only for the couple, but also for the entire family.

Support for the Mothers

Mothers described a wide range of formal and informal supports that were sought based on their main concerns and priorities, and at different times of the child’s life and of the family’s life phase.

Formal support. In the initial phase and during the diagnostic process, most mothers turned to doctors from different specialties: general practice and family medicine, pediatrics, developmental pediatrics, and child psychiatry. All mothers reported having support from a general practitioner and family doctor, a pediatrician, a developmental pediatrician and/or a child psychiatrist. All children had therapeutic support, which may include one or two of the following areas: psychology, speech therapy, and occupational therapy. The vast majority of parents became aware of these forms of support through developmental pediatricians or child psychiatrists.

All mothers reported that their children attend kindergarten at the time. The vast majority of families were supported by EI local teams. Five mothers learned about the support of the SNIPI through the kindergarten teacher, while the remaining mothers were informed by the developmental pediatrician or child psychiatrist, after the ASD diagnosis was confirmed.

Informal support. Most mothers said they could count on current family support. It is also noted, however, that this support corresponds only, in most cases, to the family nucleus (spouse and siblings), as one mother said “Support comes from the home, us three”. One mother said that she only could count on the spouse (“It is only me and my husband. Apart from us, there is no one else”), while three mothers highlighted the support from the siblings (“My daughter just naturally helps... if someone tries to speak to him and asks why he doesn’t speak”).

Taking into consideration the responses of mothers, it appears that only a few families have resorted to informal support in the initial time of diagnosis. Only two mothers emphasized the importance of the support given by other parents of children with ASD in their daily lives. One participant mentioned receiving parent to parent support.

As a result, the Internet is the informal support most used by these mothers before and during diagnosis process. First, mothers wanted to search for information and clarify doubts about the early warning signs they identified in their children. At this level, one mother highlighted “Searching on the Internet started more or less at that time, between 1 ½ and 2 years of age. I typed in some signs that my child was showing and the word Autism appeared”. After the diagnosis, mothers searched on the internet for information about the diagnosis and about the formal support available for their children. One participant highly appreciates this support today, saying “I speak to Google, my best friend (...). I read a lot of things on the Internet. I also read a lot of things from other parents who talked about what they did and this gave me some ideas”.

Early Intervention in Autism Spectrum Disorder / 245
Adequacy of the Forms of Support in Light of the Concerns of the Mothers

Formal support response. Regarding the initial response of formal support, some mothers stated that their concerns were deemphasized or devalued by doctors of the medical specialties of general practice and family medicine and pediatrics. One mother mentioned “When he was 18 months old, I spoke with a pediatrician but all he said was that my child was slightly behind but he was smart . . . there was nothing wrong with him”. Only three mothers mentioned the appreciation of the initial concerns by doctors of the medical specialties of general practice and family medicine, and pediatrics.

Three mothers also mentioned the value of the initial concerns on the part of kindergarten teachers, even for specialized medical support: “The kindergarten teacher spoke with another mother who was a nurse and who got me a consultation in the hospital because my husband’s family doctor refused to see my child or have her as his patient”.

About the adequacy of formal support in the diagnostic process, three mothers stated that the process leading to the diagnosis is complex and time consuming, and highlighted the following aspects. The waiting time for a consultation in the pediatric development specialty in a public hospital was very long: “Because we have had to wait so long for a consultation, my daughter was already three when she finally attended a pediatric development consultation . . . I have had my suspicions since she was two . . .”. The process involved in getting a consultation and carrying out exams necessary to reach a diagnosis is extremely time-consuming. These delays generate difficulties in accessing therapeutic and educational support. Some mothers mentioned that, at the beginning and during the diagnosis process, doctors need to be better informed about existing therapeutic, educational and social support. As one mother said, “They should be a little more prepared, at least with the basics and, later, help guide us in the right direction”.

Half of the families who participated in this study said that the family has to be proactive in seeking and accessing the therapeutic, educational and social support they need: “I was the one who had to make phone calls, look up on the Internet, and started asking questions because when we get the diagnosis, we are left on our own. That’s what we’ve got and each one has to survive”.

Most mothers referred to the adequacy of responses of medical specialties in the area of child psychiatry and developmental pediatrics in relation to the family’s concerns. These opinions relate mostly to the period prior to obtaining the diagnosis and to the present day, with one mother adding, “There is always someone there who is willing to help us get through this difficult time”.

In relation to the adequacy of formal support to the family’s current concerns, all participants mentioned the appropriateness of current medical and therapeutic support. One mother said, “I owe them everything, even the change in our lifestyle”. However, although they feel supported, some mothers would like to have more frequent and long therapeutic support sessions. Regarding support from the EI local team, seven mothers considered the support they received to be adequate in relation to their concerns.

Informal support response. The opinion regarding the appropriateness of informal support in different periods of the life of the child and family was different. At the beginning, three mothers said their extended families devalued the early warning signs identified by parents. One mother said, “My family thought that it was something I had invented in my head”. Currently, most participants expressed satisfaction with the support provided from the household, including just the spouse and siblings. Regarding the extended family and friends, mothers considered that these supports do not fulfill their needs. One mother said, “I’m talking about the grandparents. They accept my child the way she is, but if they could avoid being with her, they would” and “In our situation, all our friends accept him but when it comes to help, they don’t help that much”.

Discussion

Summary of Findings

Research in ASD has demonstrated the importance of early diagnosis and intervention fo-
cusing on the child and the family, and on the development of interdisciplinary answers (Gabovitch & Curtin, 2009; Reis et al., 2014). In this sense, international research has been attempting to understand the concerns of families of children with ASD and the forms of support they have access to (Hall & Graff, 2010, 2011; Siklos & Kerns, 2006). However, scientific research at the level of family centered practices in families of children with ASD is still emerging (Gabovitch & Curtin, 2009).

Findings from this investigation enhance our understanding about the perception of these families, regarding the adequacy and effectiveness of answers given by the services and supports, taking into account the concerns and priorities of the mothers of children with ASD.

Based on the qualitative data produced by this investigation, it appears that most of the concerns of the participants are expressed by most of the scientific studies in this area (Siklos & Kerns, 2006, 2007). In the initial moments, the first concerns correspond to the identification of developmental warning signs in children by the family. The absence of speech and eye contact, and the presence of unusual behavior, restricted and repetitive interests and activities compromise the process of communication between parents and the child, generating feelings of frustration and the perception of a lack of parenting competence, which is discussed in literature (Hall & Graff, 2010; Hoffman et al., 2009). After confirmation of the diagnosis of ASD, we found that the concerns increased and were no longer just child-centered, but involved the entire family and all family contexts, as is evidenced by the results by Nealy et al. (2012), and Stuart and McGrew (2009).

Data collected in this study shows that these mothers rarely resort to informal forms of support, which is contradictory to most studies in EI (Hall & Graff, 2010; Trivette, Dunst & Hamby, 2010). Also, families seem to invest more in the family nucleus, as the informal support networks are often limited to the spouse and the siblings. The low demand for broader informal support suggested by this study appears to be related to the inadequacy of informal support in relation to the concerns of mothers, since family members (grandparents, as well as other relatives and friends) have difficulty in accepting the ASD situation and are unaware of its existence, and/or avoid dealing with a child with ASD. We also found that these mothers rely very little on friends, parent groups and other parents of children with ASD, in what appears to be a voluntary protective option aimed at avoiding judgmental and prejudicial opinions from friends and family, as well as the result of the difficulty in sharing intimate information (Hall & Graff, 2011; Sullivan et al., 2012).

The inadequacy of formal and informal support means that some mothers are proactive in seeking and accessing support. Mothers attributed this need in large part to the lack of information given by the formal supports, mainly during the diagnostic process. These data provide important findings about the importance attached to the Internet by these mothers, which is the most used form of informal support to obtain information and to clear up doubts about the diagnosis of ASD and the forms of support that are available to the child, as shown in some studies (Zaidman-Zait & Jamieson, 2007). The major drawback found in this theme refers to the quantity and quality of information available on the Internet for parents who are at an early stage of acceptance and understanding of ASD, and who do not know yet how to discriminate the information correctly, which may pose a danger and raise their concerns and questions.

Furthermore, the Internet may not always have the right answers that parents seek or, on the other hand, it may provide more information than the family can handle, at an emotional level, which may have adverse effects on the parents’ perception regarding the prognosis and the future of their child and of the family itself.

Most mothers search for and give greater value to more formal support throughout the life of the child with ASD. It appears that there is a great demand for formal support, especially at the initial phase of diagnosis, and it is inversely proportional to the satisfaction levels and perceptions of the adequacy of these same supports. Such perceptions of mothers can be understood in light of the undervaluing and postponement they felt regarding their concerns from doctors. The diagnostic process is complex and time-consuming, leading to late referral and access to services and
specific supports for children with ASD and their families.

The multiple findings that emerged from the participants’ speech concerning the diagnosis of children with ASD do not reflect the directions derived from research, which stress the priority and advantages of an adequate signaling of children with ASD and early intervention, namely: (a) access to specialized information for parents, addressed to their specific concerns; (b) the possibility of preventing the most serious and complete manifestations of ASD by making use of the plasticity of the brain in the early stages, and also, behaviour changes that can be potentially accomplished by EI; and (c) the possibility of achieving gains in terms of adaptive and cognitive functioning of children with ASD that ultimately may decrease the concerns of these families (Reis et al., 2014; Zwaigenbaum et al., 2013).

Mothers reported the adequacy of recent formal supports (EI local teams and therapeutic support). Nevertheless, they criticize the amount of support, its frequency and the duration of the sessions provided. These findings show that the inadequacy of this support is primarily directed to national policies, more specifically, the way they are implemented and the quality of the procedures and measures available to children with special needs and their families in Portugal.

Ultimately, according to the Decree Law No. 281/09 of October 6 (2009), the SNIPI should be based on universal access, accountability and responsiveness of professionals and public entities, with benefits that shall be greater as earlier necessary support is given to the family and the child. The legislation shows a concern to ensure a system of interaction between families and institutions and, primarily, the health system, so that all children and families are properly identified and flagged as quickly as possible. However, professional practice shows us a different reality, where the implementation and applicability of the principles fall short of what is advocated and recommended in the theoretical and legal frameworks. Therefore, it is essential to promote cooperation and coordination between policies and practices, in order to analyze and assess the responsiveness and quality of the support that is given to families of children with ASD.

Study Limitations

Qualitative research requires that researchers have training and experience and are capable of performing a wide range of different techniques concerning qualitative methodology and semi-structured interviews. Despite the constant effort to perform procedures conferring validity, reliability and scientific rigor to this study, we are aware of the presence and influence of some limitations that will be discussed below.

Firstly, the large number of interviews and consequent volume of data made content analysis and interpretation time consuming. Beyond this reality, it is likely that some procedures related to content analysis, particularly the definition of units, as well as coding categories, may have been influenced by our personal and professional experience, as well as knowledge and beliefs about the subject.

Secondly, this study adds important considerations in terms of the adequacy of the support that is given to these families. However, we consider that there is still much to explore about this topic and, perhaps, the third research question we posed may have been too ambitious for what the study could answer. Furthermore, the 3 to 6 year old age span is, possibly, too wide, during which the concerns and the developmental and contextual demands for children with ASD change quite a bit. The selection of a more narrow age group could have been more appropriate.

Even in conceptual terms, the word “support” was not entirely clear to these families. In fact, for some participants, if a determined form of support did not meet their concerns and priorities, it was no longer considered “support”. Given the conceptual difficulty, in some interviews, it was necessary to define and clarify the designation of “support”, to ensure adjusted and detailed data, according to the interview questions.

Thirdly, we would like to mention the fact that the 12 mothers belong to a single geographic region and are mostly supported by the same entities that guarantee medical, therapeutic and educational support. This may
justify the uniformity of some of the results and experiences shared by families with regard to the adequacy of existing formal supports.

Future Research Recommendations

Considering the increasing prevalence of ASD and the insufficient amount of research in terms of a family-centered approach in the field of this disorder, we believe that there are still many paths to be discovered and created by the scientific community. This reality reinforces the need to perform further studies in this field, towards the development and implementation of strategies and practices that are increasingly tailored to the concerns and priorities of these families.

In order to achieve a broader picture of the situation of families of 3 to 6 year-old children with ASD, we recommend the development of a mixed research with a larger number of families. In this case, it would be important to include the fathers of children with ASD, to examine their perceptions about the concerns, the forms of support that families have access to, and their adequacy in light of the concerns of families. Also, the establishment of a focus group with participants would be recommended, to carry out more focused interviews, which would help to better understand the experiences and opinions of the families from an in-depth perspective.

The main findings from this study add an important insight for health professionals and the practices of early interventionists. Firstly, these results indicate an increased demand for formal support throughout the life of the child with ASD at the expense of informal support. At the same time, families refer to the inadequacy of this form of support at certain times, which is corroborated by several studies in the field of ASD (Minnes & Steiner, 2009; Renty & Roeyers, 2006). Further research is, therefore, essential to understand why current formal supports are inadequate and to cast some light on which procedures may be included, removed or altered so that the concerns, priorities and expectations of families are met. These findings are an asset to the EI professionals to: (a) have more knowledge and raise awareness about the difficulties experienced by families; (b) seek early referral to appropriate support (Siklos & Kerns, 2007); (c) assist families in building solid support networks, when such support lacks or is insufficient (Hall & Graff, 2010); and (d) assist in the coordination between the various sources of support (Renty & Roeyers, 2006).

Considering the new findings obtained in this study related to informal supports, it is relevant to research and understand parental expectations more deeply, as well as formats of search and access to, and also, the adequacy and effectiveness of, the informal support that are available to these families.

Another area for future research consideration concerns the development and concrete definition of procedures and preventive measures at the level of EI in Portugal, which are focused on raising awareness and training of professionals in primary health care (general practice and family medicine and pediatrics), and who are the first these families resort to looking for support.

A chance to put these measures and preventive strategies in practice may entail that EI local teams access more directly and quickly to the medical system. Conducting information sessions and workshops to families and doctors of primary health care, as well as regular meetings with these professionals and early interventionists, could be helpful to improve, enhance and accelerate the signaling process of these children.

It is considered that principles and procedures for the evaluation and support of families of children with ASD, even if involved and guided by a family-centered approach, present specific features which should be targeted for recognition, consideration and analysis by medical, therapeutic and educational entities, as well as by existing political entities. It seems that there is a gap between what the law states and hopes to accomplish, and what is done in the field. Therefore, evidence-based studies and further research with families and professionals must be done, in order to explore what is happening in the field of EI in Portugal, as well as the effective fulfillment of the goals advocated by the SNIPI.
References


Lima, J. A. (2013). Por uma análise de conteúdo mais fiable [For a more reliable content analysis]. *Revista Portuguesa de Pedagogia, 47*(1), 7–29. doi:10.11419/1647-8614_47-1_1


Received: 19 June 2016
Initial Acceptance: 23 August 2016
Final Acceptance: 10 October 2016
Comprehension of Verbal and Visual Metaphors among Individuals with Intellectual Disability with and without Down Syndrome

Shlomit Shnitzer-Meirovich, Bar-Ilan University and Levinsky College of Education

Hefziba Lifshitz-Vahab
Bar-Ilan University

Nira Mashal
Bar-Ilan University and Gonda Multidisciplinary Brain Research Center

Abstract: Studies examining the visual and verbal metaphorical comprehension of individuals with non-specific intellectual disability (NSID) are remarkably scarce; to date only one case study examined comprehension of metaphors in an individual with Down syndrome (DS). The current study explored both the understanding of conventional and novel metaphors and the comprehension of visual metaphors in individuals with NSID (aged 15–59, N = 53) and individuals with DS (aged 15–52, N = 50). Both etiology groups showed poor understanding of verbal and visual metaphor comprehension with worse performance on the visual task. However, the comprehension of novel metaphors was significantly higher than the comprehension of conventional ones, in both groups. As expected, individuals with DS understood fewer conventional metaphors than individuals with NSID, probably due to the linguistic deficiency characterizing individuals with DS. These findings were explained in light of the delay in linguistic ability that characterizes individuals with DS as well as the gradual expansion of their mental lexicon with increasing age.

Figurative language, such as metaphors, metonymies, simile, irony and idioms, is widely used in daily conversation (Lakoff & Johnson, 1980; Pollio, Barlow, Fine, & Pollio, 1977). The ability to comprehend and produce metaphors is critically important for communication (Gibbs, 1994; Lakoff & Johnson, 1980), and reflects the cognitive level of human creativity, the capacity for abstract reasoning and verbal abilities (Friemoth & Kamhi, 1990; Vosniadou, 1987). In order to comprehend metaphors, one must hold enough knowledge to discover a similarity between the two unrelated concepts involved (the vehicle and the target of the metaphorical expression), and to establish the common ground between them by comparison or analogy (Barcelona, 2003; Graesser, Long, & Mio, 1989; Keil, 1986; Shen, 1992; Warren, 1992). There is evidence suggesting that metaphorical comprehension is impaired in individuals with developmental disorders, including autism-spectrum disorders (Baron-Cohen, 1997; Mashal & Kasirer, 2011), learning disabilities (Cain, Oakhill, & Lemmon, 2005; Mashal & Kasirer, 2012), and Williams syndrome (Annaz et al., 2013; Van-Herwegen, Dimitriou, & Rundblad, 2013). Studies examining visual and verbal metaphorical comprehension among individuals with non-specific intellectual disability (NSID) are remarkably scarce; to date only one case study examined metaphorical comprehension in an individual with Down syndrome (DS) (Papagno & Vallar, 2001). The main aim of the current study is to explore verbal and visual metaphorical comprehension among individuals with NSID and individuals with DS.

The current study focused on two types of verbal metaphors: conventional and novel metaphors, which differ in their level of conventionality. The meanings of novel meta-
Metaphors are not coded in the mental lexicon; according to some scholars, they are understood via a comparison processes ("X is like Y") in which the semantic features of both concepts involved (the vehicle and the target) are extracted and then matched with one another (Bowdle & Gentner, 2005). The shared and non-shared elements are then used to establish the ground for the comparison. In contrast, conventional metaphors are understood via a categorization processes ("X is Y") (Glucksberg & Keysar, 1990). Accordingly, the target concept becomes a member in subordinate, abstract metaphorical category, represented by the vehicle term (Glucksberg, 2001). The "career of metaphor" model (Bowdle & Gentner, 2005) combines these two views and suggests that the manner in which a metaphor is comprehended depends on its level of conventionality. This approach postulates a shift in the mode of processing from comparison to categorization as metaphors are conventionalized. Thus, novel metaphors are primarily understood via a comparison process but after repeated use they become conventional. Once conventional, they are understood via a categorization process and their meanings can be easily retrieved from long-term memory (Bowdle & Gentner, 2005; Zharikov & Gentner, 2002).

These two modes of metaphor processing seem to utilize different cognitive abilities. Conventional metaphors are associated with rich language skills, linguistic discourse and knowledge of concepts that enable fast retrieval of their meaning as coded in the mental lexicon (Amanzio, Geminiani, Leotta, & Cappa, 2008; Gioia, 1997; Kaufman, 2001). On the other hand, the meaning of a novel metaphor is not coded in the mental lexicon; therefore, it requires formation of a novel interpretation that relies on executive function capabilities such as mental flexibility, preservation and working memory capacity (Friemoth & Kamhi, 1990; Glucksberg, 2008; Mashal & Kasirer, 2011, 2012; Silvia & Beatty, 2012; Vosniadou, 1987) as well as on analogical thinking abilities (Amanzio et al., 2008; Cameron, 1996, 2008; Gentner & Bowdle, 2008; Holyoak & Thagard, 1989; Kaufman, 2001; Kelly & Keil, 1984, Winner, 1979, 1988). Beatty and Silvia (2013) suggested that crystallized intelligence is related to conventional metaphor production, whereas fluid intelligence is associated with the production of novel metaphors. Analogical thinking is one component that comprises fluid intelligence (Lifshitz, Weiss, Tzuriel, & Tzemach, 2011). Several studies have shown a link between analogical and metaphorical thinking (Gentner, 1977; Gentner & Bowdle, 2008; Gibbs, 1994, 2008; Nippold, 1998). Gentner (1983) defines an analogy as mapping information from two concepts while maintaining an equal ratio between them (1:1). Gentner and Bowdle (2008) argued that similar to analogical processing, a metaphor is first and foremost processed symmetrically from the vehicle to the target and vice versa, and only afterwards, an inference is conducted from the vehicle to the target. Thus, analogical and metaphorical thinking establish and share three cognitive components: similarity between elements, inference between the vehicle and the target by matching the relationship between them, and identification of the target or the content to which the metaphor or the analogy refers (Holyoak & Thagard, 1989).

In addition to the two types of metaphors, the current study focused on two presentation modes, visual and verbal. The Metaphoric Triad Test (MTT) examines the ability to comprehend visual metaphors (Kogan, Connor, Gross, & Fava, 1980). In this test, the participants are presented with three pictures (e.g., a baby, a rose bud, a funnel), where two pictures are linked metaphorically (the baby and the rose bud) and one picture is a distractor (the funnel) that forms a literal link to one of the pictures. The ability to comprehend visual metaphors is demonstrated when the participants find the connection between the pair of pictures that relate metaphorically and formulate a sentence that expresses the metaphorical connection between them (e.g., a baby is a rose bud). According to Paivio’s dual coding theory (Paivio, 1991), pictures have an advantage over words since semantic information is coded via two separate rotes. Whereas words are processed only via a verbal route, pictures are processed both via an image pathway and via a verbal code. Thus, pictorial information increases the strength of encoding by accessing semantic knowledge through parallel pathways, facilitating improved recall. The ability to recall visual information
better than verbal information is known as the "picture superiority effect" (Kail & Siegel, 1977). The study by Defeyter, Russo and McPartlin (2009) demonstrated a developmental trend in the picture superiority effect in recognition memory. According to this study, while 7-year-old children with TD showed preference for words over pictures, the picture superiority effect was exhibited significantly among 9 year olds, 11 year olds and adults.

Both visual metaphors and novel metaphors require processing of items that are not coded in the mental lexicon, and therefore may rely on the recruitment of the same cognitive abilities. Mashal and Kasirer (2012) conducted a principal component analysis in order to examine the interrelationship between various cognitive abilities such as semantic knowledge, executive functions, similarities, reading fluency and the comprehension of visual and verbal metaphors. The results showed that the cognitive abilities required for comprehending visual metaphors and novel metaphors in individuals with TD are similar in relying on executive functions and analogical thinking abilities. Hence, the researchers concluded that the degree of comprehension does not necessarily depend on the metaphorical presentation modality (visual or verbal), but on the conventionality level.

Cognitive Profile of Individuals with Non-specific ID and with Down Syndrome

Individuals with ID constitute a heterogeneous group with differing IQ levels and etiologies. The current study focuses on participants with mild and moderate ID (IQ = 40–70) with NSID and with DS. Individuals with NSID are characterized by linguistic delays, such as poor language competence (Aitchison, 2003; Fink & Cegelka, 1982), limited vocabulary (Borkowski & Büchel, 1983; Cornoldi, Giofre, Orsini, & Pezzuti, 2014), lack of verbal rehearsal and reduced ability in active retrieval of coded information (Hulme & Mackenzie, 1992). They also exhibit difficulties in executive function abilities, such as dealing simultaneously with several aspects of a problem, lack of spontaneous strategy use (Borkowski, Carr, & Pressley, 1987) and difficulty in comprehending abstract relationships between pairs of objects (Paour, 1992). Individuals with NSID show inefficient short-term memory (Reed, 1996) and deficits in the working memory system (Carretti, Belacchi, & Cornoldi, 2010; Cornoldi et al., 2014; Numminen, Service, & Ruoppila, 2002; Schuchardt, Gebhardt, & Maehler, 2010). One might argue that the deficits in the working memory system, especially in the executive functions and inhibition control, might pose barriers in acquiring adequate analogical reasoning in this population (Hulme & Mackenzie, 1992; Inhelder & Piaget, 1958; Jensen, 1970; Conaghy & Kirby, 1987a).

Despite the abovementioned limitations, evidence suggests that ID is not one entity and thus individuals with the same IQ may exhibit strengths and weaknesses in different cognitive domains (Vicari, Albertini, & Caltagirone 1992). For instance, Vicari et al., (1992) examined the cognitive profile of adolescents with NSID with the same IQ and chronological age (CA) and found poorer performance in verbal ability tasks, visuo-intelligence (as measured by the Raven Matrices test, 1960) and visuo-constructive abilities (e.g. copying and drawing) than in visuo-perceptive abilities (as measured by the overlapping pictures test). These results indicated that some cognitive abilities are better preserved in individuals with NSID than others. Similar to individuals with NSID, the cognitive profile of individuals with DS is characterized by a global cognitive deficit, although some abilities are better preserved than others. Individuals with DS exhibit a remarkable deficit in language abilities relative to non-linguistic cognitive abilities, especially visual-spatial abilities (Bellugi, Bihrl, Jernigan, Trauner, & Doherty, 1990; Chapman, 1995; Fowler, 1990; Gunn & Crombie, 1996). Similar to individuals with NSID, individuals with DS exhibit difficulties in executive function skills, especially in working memory and in planning capabilities (Costanzo et al., 2013; Lanfranchi, Baddeley, Gathercole, & Vianello, 2012; Lanfranchi, Jerman, Dal Pont, Alberti, & Vianello, 2010). In addition, there is evidence of difficulties in analogical problem solving abilities (Buckley, 1985; Natsopoulos, Christou, Kiutselini, Raftopoulos, & Karefullidou, 2002).
As a result of the cognitive profiles characterizing both etiologies, one might argue that these individuals cannot go beyond a concrete level of reasoning, and would encounter difficulties in comprehending verbal and visual metaphorical language. The effectiveness of intervention programs in helping individuals with ID, with and without DS, acquire higher order cognitive abilities, such as analogical reasoning, is well documented (Büchel, Schlatter, & Scharnhorst, 1997; Hessels-Schlatter, 2002; Lifshitz, Tzuriel, & Weiss, 2005; Lifshitz et al., 2011; Conaghy & Kirby, 1987a, 1987b; Tzuriel & Klein, 1985). These studies demonstrated that adolescents and adults with mild and moderate ID can improve their level in solving conceptual and perceptual analogies as a result of a dynamic assessment procedure. Accordingly, the current study aims to investigate the ability of individuals with NSID and individuals with DS to perform higher level cognitive tasks that rely on analogical reasoning capacity, as required to comprehend metaphorical language.

The picture superiority effect was observed in individuals with NSID and individuals with DS in numerous studies on memory functions (Cherry, Applegate, & Reese, 2002; Dulaney & Ellis, 1991; Leven, Lyxell, Andersson, & Danielsson, 2013; Lifshitz-Vahab & Vakil, 2014; Simon, Rappaport & Agriesti, 1995). These studies suggest that individuals with ID can utilize nonverbal memory codes to support long-term retention as effectively as individuals with TD, and that processing visual information enables individuals with ID to develop the meaning behind visual stimuli better than words alone. Thus, it can be assumed that individuals with ID, with and without DS, will comprehend visual metaphors better than verbal ones. Moreover, the deficit in the linguistic abilities found in individuals with DS is more prominent than other populations with ID, such as individuals with Williams’s syndrome and individuals with NSID (Fowler, Gelman, & Gleitman; 1994; Hulme & Mackenzie; 1992; Mundy, Sigman, Kasari, & Yirmiya, 1988). Accordingly, it can be assumed that individuals with DS will comprehend verbal metaphors (especially conventional metaphors) less well than visual ones and less well than individuals with NSID.

Metaphor Comprehension in Individuals with ID with and Without Down Syndrome

Only few studies have examined the comprehension of metaphorical language among individuals with ID. Most of the studies focused on individuals with Williams syndrome (WS), since they possess high language abilities, compared to their spatial abilities (Jarrold, Baddeley, Hewes, & Phillips, 2001). The results of the studies indicated that participants with WS were less able to comprehend metaphorical language than their TD peers matched on CA. The WS group experienced difficulties in finding the figurative expressions in stories, avoided answering due to the lack of knowledge, and exhibited a tendency to provide more literal interpretations (Annaz et al., 2009). In a comparison between individuals with WS and individuals with TD matched by mental age (MA) and individuals with TD matched by CA, the comprehension of figurative language of participants with WS was significantly lower than that of participants with TD matched by CA, but not significantly different than the comprehension of those with TD matched by same MA (Godbee & Porter, 2013).

Only one case study examined the ability of a woman with DS (CA = 30, verbal IQ = 80, performance IQ = 63) to comprehend verbal metaphors and idioms (Papagno & Vallar, 2001). The general cognitive profile of this woman was exceptional compared to other individuals with DS. While most of individuals with DS are characterized by poor language development and delayed non-verbal cognitive abilities (Chapman, 1995), the results demonstrated reserved language abilities. On the other hand, this woman presented difficulties in executive functions, spatial perception, and comprehension of verbal metaphors and idioms; she showed a preference for giving literal interpretations to explain the metaphorical expressions presented to her. Based on her strengths and weaknesses, the authors concluded that language or literal comprehension is separate from metaphorical language comprehension and the deficiencies in comprehending metaphors and idiom are the results of deficient working memory ability, which is an important component in compre-
hending metaphorical language (Vallar & Pagano, 1993).

The aims of the current study are threefold. The first aim is to examine the differences between verbal and visual metaphorical comprehension within each etiology group. In light of the picture superiority effect (Kail & Siegel, 1977) observed among individuals with and without DS (Cherry et al., 2002; Levén et al., 2013; Simon et al., 1995), we hypothesized that both groups will perform better on the visual metaphor than on the verbal metaphorical test (conventional and novel). The second aim is to examine group differences in comprehension of metaphors on different levels of conventionality (conventional versus novel). We hypothesized that individuals with NSID will exhibit better comprehension of conventional metaphors than those with DS, due to the linguistic delay which characterizes individuals with DS. Regarding group differences in comprehending novel metaphors, we did not have a clear hypothesis since both etiology groups exhibit difficulties in executive functions and analogical abilities, which are associated with the ability to comprehend novel metaphors (Mashal & Kasirer, 2012). The third aim is to test the relationship between metaphorical comprehension (verbal and visual) and three fundamental cognitive abilities required for metaphorical comprehension: linguistic, executive functions, and analogical abilities. We predicted positive correlations between linguistic, executive functions and analogical abilities and comprehending metaphorical language.

Method

Participants

Our sample was comprised of 103 participants with mild and moderate ID (IQ = 40–70) divided into two etiology groups: 53 participants with NSID (30 boys and 23 girls, CA = 15–59) and 50 participants with DS (27 boys and 23 girls, CA = 15–52). The two etiology groups did not differ in gender $\chi^2 (1) = .07, p = .79$. The linguistic ability of the participants was assessed using the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997), idioms and synonyms. The analogical ability was assessed by using Raven’s standard progressive matrices (Raven, 1960), and these results were used to compare the MA and basic cognitive level of the two etiology groups.

The PPVT has been used to match participants with ID, and measure their vocabulary knowledge (Facon & Facon-Bollengier, 1999). The test consists of 204 items. In each item, participants were presented with four pictures and asked to point to only one picture that represents the word read to them. The score is the total number of correct answers. The discontinuation criterion was six incorrect responses out of eight in the same level. The MA of the participants was calculated according to the test criteria. The Raven test was designed to assess the ability to form comparisons, deduce relationships, correlates, and reason by analogy (Raven, Court, & Raven, 1986). The participants solved sets A, B, C, D and E. In each set, the test was stopped after five consecutive incorrect responses. Correct answers received 1 point. The scores were produced by summarization of the raw scores.

The two etiology groups were also matched according to CA. The differences between groups on the CA, MA and the Raven tests were analyzed by T tests for independent samples. As can be seen in Table 1, the NSID group did not differ significantly from the DS group in these three variables.

Participants were recruited from schools, residential and vocational facilities of adolescents and adults with ID under the supervision of the Division of Chief Scientist of Israel Ministry of Education and the Division of Intellectual Disability of Israel Ministry of Welfare. All of the participants met the criteria set for the current research: individuals with mild/moderate ID according to the traditional

<table>
<thead>
<tr>
<th></th>
<th>NSID (N = 53)</th>
<th>DS (N = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean SD</td>
<td>Mean SD</td>
<td>t(100)</td>
</tr>
<tr>
<td>CA</td>
<td>28.49 12.20</td>
<td>27.56 11.43</td>
</tr>
<tr>
<td>MA</td>
<td>9.79 1.97</td>
<td>9.28 2.08</td>
</tr>
<tr>
<td>Raven</td>
<td>12.91 3.44</td>
<td>13.72 3.61</td>
</tr>
</tbody>
</table>
AAMR definition (Grossman & Begab, 1983), independent in terms of Activities of Daily Living (ADL) skills and without maladaptive behavior.

**Material**

The participant’s comprehension of metaphors and executive function abilities were assessed using the phonemic and semantic fluency tests: the Metaphoric Triad Test (MTT), a metaphoric questionnaire, the Trail Making Test (TMT), and the Homophone Meaning Generation Test (HMGT).

**Metaphoric Triad Test (MTT)**

The MTT examines the comprehension of visual metaphors (Kogan et al., 1980). We used 20 pairs of pictures that have a metaphorical relationship (a lady with closed eyes, shutters). The participants received 1 point if they formulated the relationship between the two pictures as simile (“the eyes of the lady are closed like shutters”) or as a metaphor (“closed eyes are shutters”) for each pair of pictures. A literal relationship (“the women is sleeping in her home”) received 0 points (range 0–20). Prior to the test, the participants received three examples of paired pictures. The experimenter pointed out the literal and metaphorical relationship for each example pair of pictures in a random order, so the participants could not guess which interpretation is preferable.

**Verbal Metaphors**

The metaphoric questionnaire tested the participant’s ability to comprehend two types of metaphors: 10 conventional (e.g., sharp tongue) and 10 novel (e.g., pure hand) (Mashal & Kasirer, 2011). For each of the metaphorical expressions, four interpretations were offered: a correct metaphorical interpretation, a literal distracter, an unrelated interpretation, and a fourth choice, “this expression is meaningless.” The metaphorical expressions and the four interpretations were read to the participants. A correct answer received 1 point (range 0–10 for each type of metaphor). Furthermore, the participant’s linguistic ability was assessed by idioms and synonyms tests.

**Idioms**

This test examines the comprehension of idioms (20 items, e.g., “he got cold feet”) (Mashal & Kasirer, 2011). For each, four interpretations were offered: a correct idiomatic interpretation (“he lost courage”); a literal interpretation (“the temperature of his feet lowered”); a literal distracter (“he got a present”); an unrelated interpretation (“he explained himself”). The idiomatic expressions and the four interpretations were read to the participants. A correct answer received 1 point (range 0–20).

**Synonyms**

The synonyms test is one of the nine scales of the MANN Hebrew test (Glanz, 1989) for vocabulary. The participants were presented with 12 key words and asked to identify a similar word for each, on a list of five other words that were read to them (e.g., “The synonym of the word ‘wall’ is: gate, path, way, balcony or side.”). A correct answer received 1 point (range 0–12).

Participants’ executive functions abilities were assessed using the phonemic and semantic fluency tests, TMT and HMGT.

**Phonemic fluency test.** The phonemic fluency test examines verbal knowledge, flexibility, and executive control (Kave et al., 2010). The participants were asked to provide as many words as possible beginning with each of three letters (b, g, sh) within 60 seconds. The score is the sum of the words generated for the three letters.

**Semantic fluency.** The semantic fluency test examines flexibility, and executive control (Kave et al., 2010). The participants were asked to provide as many words as possible in three semantic categories (animals, fruits and vegetables, and vehicles) within 60 seconds. The score is the sum of the words generated for the three categories.

**Trail Making Test**

The TMT (Reitan & Davison, 1974) examines executive function processes (Espy & Cwik,
2004), and is widely used in neuropsychology studies (Lezak, Howieson, & Loring, 2004). Part A measures visual scanning and tracking, motor speed and focused attention. The participant is asked to draw lines that connect consecutive digits printed in a scattered pattern on the page. Part B measures cognitive flexibility, set shifting and divided attention. The participant is asked to draw lines that connect sequences of letters and digits, alternatively e.g., 1-A-2-B and so on). The total score is the time in seconds needed to complete the two tests. The test was conducted in a population of ID by Smith (1963).

**Homophone Meaning Generation Test**

The HMGT (Mashal & Kasirer, 2011) examines the ability to shift between the different meanings of a homophone (10 items in Hebrew). The participants were presented with 10 sentences in which the final word ended with a homophone (e.g., “He wrote a letter”). The participants were instructed to say aloud all of the meanings of the homophone. A correct answer received 1 point.

**Procedure**

Consent for participation in the study was obtained from the parents or guardians of the participants. Authorizations were obtained from the Ministry of Education, Ministry of Welfare and the University Ethics Committee. The aim and procedure of the study were explained to all participants. All of the participants agreed to participate in the study. They signed an informed consent form for participation in the study. In line with the “normalization principle” (Wolfensberger, 2002), the participants received payment or a gift, according to their choice.

The study was carried out in two stages. In the first stage, each participant was assessed individually using the PPVT and the Raven tests. This stage lasted 2 hours. The tests were administered in a quiet room in the participants’ vocational and residential facilities. The second stage was conducted several days later and was divided into two parts: in the first session, the verbal metaphoric test and the MTT (in a PowerPoint presentation) were administrated along with the idioms, HMGT, semantic and the phonemic fluency test. After a 45-minute refreshment break, the synonym and the TMT were administered. This stage lasted two to three hours. The tests were administered to the participants in a fixed order.

**Results**

**Metaphorical Comprehension**

To examine performance on the visual and two verbal metaphor (conventional and novel) tests, the participants’ performance on the three tests was first converted to percentages. Next, we performed a 2×3 repeated measures ANOVA, with group (NSID, DS) as the between-subject variable and the type of metaphor (conventional, novel and visual metaphors) as the within-subject variable.

No main effect of group was found, $F(1,101) = 1.09, p = .30, \eta^2_p = .01$. The main effect of the type of metaphor was significant, $F(2,100) = 107.72, p < .001, \eta^2_p = .68$. Bonferroni post hoc analysis indicated that the performance on the visual metaphor test ($M = 7.18\%, SD = 9.64$) was significantly lower than the performance on the novel ($M = 30.29\%, SD = 16.65$), $p < .001$, and the conventional ($M = 19.61\%, SD = 11.37$), $p < .001$ metaphor tests. Furthermore, performance on the novel metaphors test was significantly higher than on the conventional and visual ones, $p < .001$. Moreover, the group X type of metaphor interaction was significant, $F(2,100) = 5.56, p < .01, \eta^2_p = 0.10$. Bonferroni post hoc analysis further revealed that the source of the interaction is the higher performance on the conventional metaphors test in the NSID group than in the DS group, $p < .01$. No significant group difference was found for performance on the visual ($p = .18$) and novel metaphor tests ($p = .68$) (see Figure 1).

**Prediction of Metaphor Comprehension Scores by Crystallized and Fluid Scores**

To examine which of the three fundamental cognitive abilities (linguistic, executive functions, and analogical thinking) predict the comprehension of visual and verbal metaphors, we used a set of six hierarchical and stepwise regression analyses, three for each etiology group. We conducted one analysis...
with performance on the visual metaphors test as the predicted variable, another analysis using performance on the conventional metaphors test, and a final one using performance on the novel metaphor test. In each regression, we first entered the scores of the background characteristics, PPVT, Raven and CA. Next, we entered the scores on the linguistic tests (idioms and synonyms) and the executive functions tests (fluency tests, TMT and HMGT). The results of the three regression analysis in the NSID group are presented in Table 2 and the results of the three regression analysis in the DS group are presented in Table 3. The order of the variables presented the order of significance.

As can be seen in Table 2, the PPVT scores explained a significant share of the variance in performance on the novel and visual metaphors, but not in performance on the conventional metaphor test.

As can be seen in Table 3, the PPVT scores explained a significant share of the variance in performance on the novel metaphor test, but not on the visual and conventional metaphor tests. Scores on executive function abilities, measured by the phonemic fluency test, added significantly to the share of variance in the performance on the conventional and visual metaphors tests. Moreover, the HMGT added significantly to the share of variance in the performance on the novel metaphor test, but not on the visual and conventional metaphor tests. The CA contributed to the share of variance only for the performance on the conventional metaphors test.

**Discussion**

There are three main issues at the core of the discussion. The first issue relates to the differences between verbal and visual metaphorical comprehension within each etiology group. The second issue concerns the differences between the two etiology groups in verbal metaphorical comprehension in respect to the con-

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conventional metaphors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>.13</td>
<td>.02</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.00</td>
<td>.01</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>−.03</td>
<td>.05</td>
<td>−.07</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td><strong>Novel metaphors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>.00</td>
<td>.02</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.04</td>
<td>.01</td>
<td>.42**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.02</td>
<td>.07</td>
<td>.05</td>
<td>.19*</td>
<td></td>
</tr>
<tr>
<td><strong>Visual metaphors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>.03</td>
<td>.09</td>
<td>.05</td>
<td>.346***</td>
<td>.346***</td>
</tr>
<tr>
<td>PPVT</td>
<td>.27</td>
<td>.06</td>
<td>.58***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.15</td>
<td>.31</td>
<td>.06</td>
<td>.346***</td>
<td>.346***</td>
</tr>
</tbody>
</table>

**p < .01, ***p < .001.**
ventionality level (conventional versus novel metaphors). The third issue relates to the contribution of linguistic, executive functions and analogical abilities to predicting the comprehension of visual and verbal metaphors.

### Verbal and Visual Metaphorical Comprehension of Individuals with NSID and with DS

The main aim of the current study was to explore verbal and visual metaphorical comprehension of individuals with NSID and individuals with DS. In general, the results indicated that the overall performances on the verbal and visual metaphorical comprehension tests in both etiologies were very poor. There are two possible explanations for this finding. First, the cognitive profiles that include delayed linguistic development (Chapman, 1995; Cornoldi et al., 2014; Fowler, 1990), executive functions (Borkowski et al., 1987; Buckley, 1985; Carretti et al., 2010; Costanzo et al., 2013; Hulme & Mackenzie, 1992) and analogical performance (Buckley, 1985; Cornoldi et al., 2014; Fowler, 1990).

### Table 3

Summary of Regression Analysis Predicting Comprehension by CA, Basic Cognitive Level Tests and Scores on Linguistic, Executive Functions and Analogical Tests in the DS Group (N = 50)

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>Blocks</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
<th>R²</th>
<th>Δ R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional metaphors</td>
<td>“Block 1”</td>
<td>.02</td>
<td>.01</td>
<td>.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.00</td>
<td>.01</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.01</td>
<td>.04</td>
<td>.03</td>
<td>.09</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>“Block 2”</td>
<td></td>
<td>.02</td>
<td>.01</td>
<td>.27*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.01</td>
<td>.01</td>
<td>.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.02</td>
<td>.04</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonetic fluency</td>
<td>.05</td>
<td>.02</td>
<td>.35*</td>
<td>.18*</td>
<td>.09*</td>
<td></td>
</tr>
<tr>
<td>Novel metaphors</td>
<td>“Block 1”</td>
<td>.02</td>
<td>.02</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.06</td>
<td>.01</td>
<td>.68***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>-.09</td>
<td>.06</td>
<td>-.21</td>
<td>.39***</td>
<td>.39***</td>
<td></td>
</tr>
<tr>
<td>“Block 2”</td>
<td></td>
<td>.02</td>
<td>.02</td>
<td>.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.05</td>
<td>.01</td>
<td>.60***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.15</td>
<td>.06</td>
<td>.35*</td>
<td>.18*</td>
<td>.09*</td>
<td></td>
</tr>
<tr>
<td>HMGT</td>
<td>.22</td>
<td>.11</td>
<td>.30*</td>
<td>.44***</td>
<td>.05*</td>
<td></td>
</tr>
<tr>
<td>Visual metaphors</td>
<td>“Block 1”</td>
<td>-.02</td>
<td>.07</td>
<td>-.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.18</td>
<td>.05</td>
<td>.49***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.37</td>
<td>.26</td>
<td>.20</td>
<td>.38***</td>
<td>.38***</td>
<td></td>
</tr>
<tr>
<td>“Block 2”</td>
<td></td>
<td>-.04</td>
<td>.07</td>
<td>-.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.15</td>
<td>.05</td>
<td>.41**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.09</td>
<td>.27</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMGT</td>
<td>1.06</td>
<td>.46</td>
<td>.35*</td>
<td>.45***</td>
<td>.07*</td>
<td></td>
</tr>
<tr>
<td>“Block 3”</td>
<td></td>
<td>-.04</td>
<td>.06</td>
<td>-.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.17</td>
<td>.05</td>
<td>.47***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven</td>
<td>.09</td>
<td>.26</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMGT</td>
<td>1.33</td>
<td>.45</td>
<td>.42***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonetic fluency</td>
<td>.28</td>
<td>.13</td>
<td>.27*</td>
<td>.51***</td>
<td>.06*</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05,  ** p < .01,  *** p < .001.
characterizing both etiology groups, as well as their difficulties in dealing with abstract cognitive problems, may lead to the observed difficulties in comprehending verbal and visual metaphorical language. Second, the poor performance could also result from the fact that individuals with ID rarely use metaphors in everyday life. Indeed, the use of simple sentence structures and concrete phrases rather than abstract ones is well documented in individuals with ID (Borkowski & Büchel, 1983; Conway & Pisoni, 2008; Miolo, Chapman, & Sindberg, 2005).

Contrary to our hypothesis, we found lower comprehension of visual metaphors than verbal metaphors. In light of the picture superiority effect (Kail & Siegel, 1977), we hypothesized that the opposite would be found. There are two possible explanations for this finding. One possible explanation for the lower performance on the visual metaphors test, is that the verbal ones may be related to the MA of the participants. The mean MA in the NSID group was 9.79 and 9.28 in the DS group. The study by Defeyter et al. (2009) demonstrated a developmental trend in the picture superiority effect. According to this study, the picture superiority effect was exhibited among individuals with TD, approximately from the age of 9 through adulthood. The MA of the participants in the current study is close to the lower limit of the MA that is expected to show the picture superiority effect. A second possible explanation may be related to the differences in task requirements. While the verbal tests asked the participants to choose between four choices in the visual metaphors test, the participants had to generate their own metaphorical utterance, based on the pairs of images presented to them. Forming the metaphorical utterance is a demanding task for individuals with ID seeing as they use figurative language infrequently in their daily conversations, preferring a simple, concrete sentence structure over an abstract one (Borkowski & Büchel, 1983; Conway & Pisoni, 2008; Miolo et al., 2005). Furthermore, while in the verbal tests, the participants had to select the correct metaphorical utterance among four choices, in the visual test, the participants also had to ignore irrelevant details related to the main object (e.g., clothes, hair, colors) and to the background, when attempting to generate the metaphorical relation between the two pictures. Intact executive function abilities, particularly inhibition control, are required to accomplish this. Both executive functions and analogical thinking abilities are delayed among individuals with NSID (Hulme & Mackenzie, 1992) and among individuals with DS (Buckley, 1985).

Conventional and Novel Metaphorical Comprehension among Individuals with NSID and with DS

As we hypothesized, individuals with NSID were better at comprehension of conventional metaphors than individuals with DS. The comprehension of conventional metaphors is associated with crystallized intelligence and is based on meaning retrieval from the mental lexicon (Amanzio et al., 2008). Individuals with DS are characterized by having a linguistic deficiency (Chapman, 1995; Fowler, 1990) that impairs the comprehension of conventional metaphors.

Regarding the conventionality level of the metaphorical utterances, the results indicated that the comprehension of novel metaphors was significantly better than the conventional ones in both groups. This finding is surprising due to the higher cognitive demands required for the processing of novel metaphors compared to conventional ones. The comprehension of novel metaphors requires higher comparative ability that places higher demand on working memory, while the comprehension of conventional metaphors is an easier task that relies on the automatic process of meaning retrieval (Mashal, 2013). One possible explanation for this surprising finding may be the fact that the meanings of conventional metaphors, same as novel ones, are not coded in the mental lexicon of individuals with ID, therefore, conventional metaphors hold no advantage over novel ones. Due to the rare use of metaphors in populations with ID, all metaphorical utterances are perceived as novel. However, the reason that the participants understood fewer conventional metaphors than novel metaphors, even if they perceived them as unfamiliar, is unclear.

One possible explanation may be related to the level of the metaphorical “opacity.” Metaphors can be characterized by their level of opacity (Levorato & Cacciari, 2002). The con-
ceptual relation between the vehicle and the target terms is easily inferable in “transparent” metaphors but in “opaque” metaphors, the conceptual relation between the two terms is not so easily inferable and more complex mapping is required. Comprehending transparent metaphors involves an inference process in which the conceptual similarity between the two terms is clear, while comprehending opaque metaphors requires the retrieval of their meaning as stored in long term memory. It is possible that the participants were better able to comprehend the novel metaphors since they were more transparent than conventional ones. For instance, the novel metaphor “safe secrets” could be more transparent and easily comprehensible to a person who has kept secrets within himself, much like a safe, whereas the conventional metaphor “deafening silence,” is more opaque, abstract and difficult to comprehend.

Another possible explanation for the better comprehension of novel metaphors in individuals with DS may be related to the atypical brain lateralization observed in this etiology. There is evidence suggesting a link between novel metaphor comprehension and increased activity in the right hemisphere language brain regions (in temporal regions homology of the classical language comprehension area, Wernicke’s area) (Faust & Mashal, 2007; Mashal, Faust, Hendler, & Jung-Beeman, 2007). Conversely, the comprehension of conventional metaphors and literal expressions is associated with increased activity in the classical language (Broca and Wernicke) areas of the left hemisphere. Many studies examining brain functions and lateralization in language tasks have shown that there is atypical brain lateralization in individuals with DS, compared to individuals with TD. For instance, while individuals with TD exhibit right ear advantage during the processing of linguistic stimuli, studies among DS showed a left ear advantage (Chua, Weeks, & Elliot, 1996; Elliott & Weeks, 1993; Grouios, Ypsilanti, & Koidou, 2013; Heath & Elliot, 1999; Menghini, Costanzo, & Vicari, 2011; Paquette, Bourassa, & Peretz, 1996). The left ear advantage indicates that the right, rather than the left, hemisphere is dominant for language processing among individuals with DS. It is possible that the right hemisphere advantage led to better performance on the novel metaphor test among individuals with DS. It could be assumed that this possible explanation is true even for participants with NSID, who presented similar findings. However, further studies exploring hemispheric lateralization and its relationship to metaphor comprehension among individuals with NSID and DS are required.

The Contribution of Linguistic, Executive Functions and Analogical Abilities to the Prediction of the Comprehension of Verbal and Visual Metaphors

The results of the regression analyses showed a significant contribution of vocabulary (PPVT) to the prediction of the comprehension of novel and visual metaphors (but not of the conventional ones) in both etiology groups. These findings are consistent with previous studies, pointing to the importance of verbal abilities to metaphor comprehension (Friemoth & Kamhi, 1990; Vosniadou, 1987). In the DS group, in addition to the contribution of vocabulary, executive functioning (as assessed by the HMGT) contributed to explaining the variance of the novel metaphor comprehension. Numerous studies found a connection between the executive functions and the ability to comprehend metaphorical expressions among participants with TD (Friemoth & Kamhi, 1990; Glucksberg, 2008; Mashal & Kasirer, 2011; Silvia & Beaty, 2012; Vosniadou, 1987). It is possible that the additional contribution of executive functions abilities to explaining the variance of the performance in the novel metaphors test occurs due to the linguistic delay that characterizes individuals with DS. This delay might lead the participants with DS to rely on their other capabilities in order to find the relationship between the pairs of images in the visual metaphors test.

Executive functions (HMGT, phonemic fluency test) also contributed to the prediction of visual metaphor comprehension among the DS group alone. This finding corroborates the results of the study by Mashal and Kasirer (2012). According to this study, comprehending novel and visual metaphors relies on executive function abilities required for the formulation of their interpretations. The HMGT examines the ability to shift between different
meanings of a homophone, and relies mainly on mental flexibility, a component of executive function (Mashal & Kasirer, 2011). Mental flexibility enables the selection of the relevant attributes of the vehicle and the target term and the switching between the literal and the metaphorical interpretations of a novel metaphor (Friemoth & Kamhi, 1990; Glucksberg, 2008; Mashal & Kasirer, 2011, 2012; Silvia & Beaty, 2012; Vosniadou, 1987). Similarly, understanding a visual metaphor requires the selection of relevant features of the two pictures and the creation of a metaphorical interpretation.

Non-verbal intelligence (as assessed by Raven) was also found to make a significant contribution to predicting novel metaphor comprehension in the DS group. The Raven matrices tests fluid intelligence and assesses analogical reasoning. Studies showed a correlation between the fluid and analogical abilities and the ability to comprehend metaphorical expressions (Gentner, 1977; Gentner & Bowdle, 2008; Gibbs, 1994, 2008; Nippold, 1998; Silvia & Beaty, 2012; Yosef, 2011). It can be assumed that the delay in language ability among individuals with DS, led the participants to rely not only on their linguistic ability but also on their executive functions and analogical reasoning abilities in order to find the relevant attributes of the vehicle and the target terms in the novel metaphors test.

The factors contributing to predicting the comprehension of conventional metaphors were different. Our results showed a significant contribution of both CA and performance on the phonemic fluency test in the DS group but not in the NSID group. The comprehension of conventional metaphors increases along with the participant’s age (Gardner, 1974; Nippold, 1998). It can be assumed that the contribution of CA in the DS group, who exhibited a remarkable deficit in language abilities, may indicate gradual expansion of the mental lexicon in this population with increasing age. The gradual expansion of the mental lexicon facilitates fast retrieval of the meaning of conventional metaphors (Giora, 1997). Moreover, there is evidence that the recognition of conventional metaphors by adults with TD is mostly based on phonological information maintained in the working memory (Mashal, 2013). The results of our study support these findings and show a significant contribution of the phonemic ability to the prediction of conventional metaphor comprehension, even in individuals with DS.

Taken together, the results of the current study suggest that individuals with ID are impaired in verbal and visual metaphor comprehension. Contrary to our expectation, based on the “picture superiority effect,” individuals with and without DS are impaired mostly in visual metaphor comprehension. Further studies are required in order to examine the effect of intervention programs on acquiring visual metaphorical comprehension among these populations. Regarding the conventionality level of the metaphorical utterances, the results indicated that the comprehension of novel metaphors was significantly better than conventional ones across groups. Our results also indicated that individuals with DS rely on different cognitive abilities rather than individuals with NSID, when interpreting metaphorical language. Linguistic ability (PPVT) contributed to novel metaphor comprehension in both groups, but analogical reasoning ability and mental flexibility also contributed to novel metaphor comprehension in the DS group. Moreover, the comprehension of conventional metaphors among the DS group was significantly lower than among the NSID group. The DS but not the NSID group demonstrated that better comprehension of conventional metaphors is linked to greater chronological age. We explained these findings in light of the delay in linguistic ability which characterizes individuals with DS and the gradual expansion of the mental lexicon with increasing age. Our findings may have educational implications and indicate the need for interventions for individuals with and without DS who exhibit a remarkable deficit in metaphor comprehension, particularly for visual metaphors.

References


Disability Research, 43, 489–496. doi:10.1046/j.1365-2788.1999.00224.x


The school-age and adolescent years (2nd ed.). Austin, TX: PRO-ED.

Received: 20 April 2016
Initial Acceptance: 27 June 2016
Final Acceptance: 18 August 2016
Functional Living Skills and Adolescents and Adults with Autism Spectrum Disorder: A Meta-Analysis

Ee Rea Hong
University of Tsukuba

John L. Davis
University of Utah

Leslie Neely
The University of Texas at San Antonio

Jennifer B. Ganz and Kristi Morin
Texas A&M University

Jennifer Ninci
University of Hawaii at Manoa

Margot B. Boles
Texas A&M University

Abstract: Functional living skills are skills needed for being an independent individual in society. As individuals with autism spectrum disorder (ASD) get older, the discrepancy between functional living skills of themselves and their peers increases. However, it is not known which type of intervention is more or less effective specifically for adolescent- and adult-aged persons with ASD. This systematic review and meta-analysis analyzed peer-reviewed research concerning functional living skills for individuals with ASD. Using the Tau-U effect size analysis, the following categories were analyzed: participant diagnoses, independent variables, and dependent variables. In addition, to identify statistically significant differences based on categories of the evaluated variables, we conducted the Kruskal-Wallis analysis and a Dunn post-hoc test. A total of 32 single-case studies were included in this analysis. Results indicated that interventions to improve functional living skills with adolescent- and adult-aged persons with ASD had overall strong effects. Moderate to strong effects were noted across categories for diagnosis. Findings indicated strong effects across categories for dependent and independent variables. Limitations and implications for practice and future research were discussed.

Autism spectrum disorder (ASD) is a life-long developmental disability that affects an estimated 1 in 68 children in the United States (Centers for Disease Control and Prevention [CDC], 2014). However, this statistic is not just limited to the United States; similar increases in ASD prevalence have been reported worldwide (see Bölte, Poustka, & Holtmann, 2008; Honda, Shimizu, & Rutter, 2005; Sun & Allisson, 2010). Core characteristics of ASD include deficits in communication and social interactions, as well as the presence of stereotypic and repetitive behaviors (American Psychiatric Association [APA], 2013). In addition, autism is often correlated with cognitive deficits and delays in adaptive functioning (Volkmar, Sparrow, Goudreau, & Cicchetti, 1987). Given deficits in several skill domains, adolescents and adults with ASD often experience negative outcomes, such as low rates of employment and low social involvement, compared to those without autism (Howlin, Goode, Hutton, & Rutter, 2004).

Functional living skills are defined as skills needed for being an independent individual in society (Sparrow, Cichetti, & Balla, 2005). As individuals with ASD get older, the discrepancy between the functional living skills of...
themselves and their peers increases (Carter et al., 1998). Given the lack of these skills, individuals with ASD often tend to be dependent on their families or related services when they reach adulthood, resulting in an increased burden of care for their families or service providers (Howlin et al., 2004). For example, Shattuck and colleagues (2012) found that approximately 80% of young adults with ASD lived at home, and among those individuals, only 6% had paid jobs after graduating from public schools. With the growing number of individuals with ASD who transition into adulthood and employment, more attention has recently been given to understanding and treating adolescent- and adult-aged persons with ASD (Cимера & Cowan, 2009); however, there remains a dearth of research on promoting the independent living of those individuals (Shattuck et al., 2012).

To improve functional living skills of individuals with ASD, researchers have examined various types of instructional approaches (e.g., prompting, reinforcement, modeling) and found them to be effective in promoting those skills (Bennett & Dukes, 2014; Walsh et al., 2014). Ninci and colleagues (2015) found that there are four types of interventional approaches that have been most utilized to improve functional living skills of individuals with ASD. Those approaches include video modeling, behavioral in-vivo procedures, audio cueing, and visual cues (Ninci et al., 2015). Moderate to strong effects were found for the four intervention techniques. Although those types of strategies were found to be effective in improving functional living skills of individuals with ASD, it is not known which type of intervention is more or less effective specifically for adolescent- and adult-aged persons with ASD because prior work was not limited to this age range.

A recent meta-analytic review (Ninci et al., 2015) indicated that most studies that investigated within-person changes in functional living skills for individuals with ASD have focused on adolescent- and adult-aged persons. Additionally, adolescent- and adult-aged participants with ASD showed strong effects compared to the preschool- and secondary-aged participants with ASD. However, it is unknown why those older participants show relatively stronger improvements in targeted functional living skills. Many researchers have investigated how IQ affects acquisition of functional living skills in individuals with ASD. These studies have found that children with ASD who had higher IQs showed better acquisition of targeted skills than children who had lower IQs, indicating that the presence of an intellectual disability may slow the skill acquisition rate (Freeman, Del’Homme, Guthrie, & Zhang, 1999; Green & Carter, 2011). Although previous studies have suggested that cognitive functioning levels may differentially affect the rate of growth among individuals with ASD, most of the studies examined its relation with early-aged individuals, leading to difficulties generalizing their findings to older individuals with ASD.

In addition to clarifying under what circumstances each intervention is effective for adolescents and adults with ASD, an evaluation of outcome variables that the intervention strategy can be expected to have a positive impact on is also needed (Horner et al., 2005). In addition, it is necessary to assess the efficacy of the intervention strategy for particular functional living skills in order to assist practitioners in choosing the most appropriate functional living skill intervention. Four functional living skills have often been targeted as outcome variables in the research literature, including self-help skills, household chores, employment skills, and skills related to accessing the community (see Ninci et al., 2015); however, no previous analyses have evaluated whether older individuals with ASD show better acquisition of a particular functional living skill compared to others, leading to difficulties for practitioners who wish to select an appropriate target behavior for this population.

Recent legislative mandates in the U.S., including the Individuals with Disabilities Education Act (IDEA) of 2004, require educators to use scientifically validated practices, also called evidence-based practices (EBPs; Horner et al., 2004). Because single-case research is the primary design methodology used for individuals with low-incidence disabilities, such as ASD and significant developmental delay, this methodology is considered appropriate for establishing EBPs with this population (Horner et al., 2004). However, single-case studies often use a variety of research designs and outcome measures, making it challenging
to aggregate and analyze data from those studies (Ganz et al., 2011). To help identify EBPs, meta-analytic techniques have been used to synthesize and analyze data from different single-case studies using a single statistic metric (Banda & Therrien, 2008).

Tau-U, a non-parametric effect size measure, has been considered the most appropriate statistical analysis for single-case research because it addresses several issues related to previously used non-overlap statistical analyses (Parker et al. 2011b). First, Tau-U offers greater statistical power and precision than other non-overlap statistic options (Parker et al. 2011b). Second, Tau-U is robust to autocorrelation of data; therefore, Tau-U does not vary in response to a level of autocorrelation of data (Parker et al., 2011b). Third, Tau-U has a function of controlling an undesirable baseline trend (Parker et al., 2011b). Fourth, as a top-down approach, Tau-U can be calculated with only a few data points and phases in the design (Parker & Vannest, 2012). In addition, Tau-U has been found to be consistent with visual analysis of data in a design (Brossart, Vannest, Davis, & Patience, 2014; Parker et al., 2011b). Tau-U has increasingly been used in single-case research studies (Ganz et al., 2013; Hong, Ganz, Gilliland, & Ninci, 2014) and meta-analytic reviews of single-case research (Bowman-Perrott et al., 2013; Ninci et al., 2015).

This meta-analysis aimed to determine the magnitudes of effect of educational interventions for teaching functional living skills to adolescent- and adult-aged individuals with ASD, differentiated by the following potential variables: (a) independent variables, (b) participant diagnoses, and (c) dependent variables. In addition, we investigated whether any statistically significant differences existed within levels or categories of the evaluated variables.

Research Questions

1. What are the magnitudes of effect (i.e., Tau-U effect sizes) of educational interventions for teaching functional living skills to people with ASD, differentiated by the following variables: (a) independent variables, (b) participant diagnoses, and (c) dependent variables?

2. Are there any statistically significant differences within levels or categories of the evaluated variables?

Method

A subset of the articles identified by Ninci et al. (2015) were also included in this review and the search and inclusion/exclusion procedures herein reflect those conducted in the prior meta-analysis. In the review by Ninci et al. (2015), studies that evaluated effects of educational interventions on improving functional living skills of all aged-individuals with ASD were included, while this review targets studies that included young adolescent- and adult-aged individuals with ASD. The studies selected for inclusion in the current meta-analysis were identified following all inclusion and exclusion procedures for Ninci et al. (2015). The search and inclusion/exclusion procedures for Ninci et al. (2015) are summarized below for the readers’ information.

Article Identification

To identify studies to include in the prior review, the following electronic databases were searched: ERIC, Academic Search Complete, Professional Development Collection, Social Sciences Full Text, and PsycINFO. Only peer-reviewed journal articles were included in this review. Publication year was not restricted. If an article did not have an author or was not a peer-reviewed article, the article was excluded. This resulted in a total of 1761 articles that were reviewed to determine whether they met the inclusion criteria.

Inclusion and exclusion criteria. To be included in the review, an article had to meet the following inclusion criteria: (a) included at least one participant with ASD (i.e., the participant was described as having a pervasive developmental disorder, Asperger syndrome, ASD or “autism,” or an “autistic” behavior); (b) used a single-case research design (e.g., multiple-baseline design, multiple-probe design, alternating treatment design, reversal or withdrawal designs, or a combination of more than one of those designs); (c) presented data in a line graph with individual data points; (d) included at least one dependent variable that targeted an independent adaptive or func-
tional living skill of an individual with ASD (e.g., house-keeping tasks, employment, transportation use, cooking, hygiene and personal care, shopping, accessing public setting, banking and money management, self-feeding, and toileting initiations); (e) included a type of educational intervention as the independent variable; and (f) was published in English. Studies were excluded if any of the following exclusion criteria were met: (a) dependent variable targeted social, play, communication, or leisure skills, (b) data on the participants’ behavior were not separated from the implementer’s behavior (e.g., level of prompt provided); or (c) used any type of pharmacological treatments.

As a result of the initial screening, a total of 44 studies met the initial inclusion criteria. To identify additional articles that might meet the inclusion criteria, an ancestral search was conducted, resulting in 28 studies. Therefore, a total of 71 articles were identified for further evaluation. Among the 72 articles, 32 articles included young adolescent- or adult-aged persons with ASD and were evaluated for further analysis in the current meta-analysis.

Inter-rater reliability on inclusion and exclusion criteria. To determine whether the study met the initial inclusion criteria for the Ninci et al. (2015) meta-analysis, two independent coders evaluated a total of 881 of the 1761 articles (50%). A percentage of agreement was used to calculate an inter-rater reliability (IRR) score throughout the study. IRR was calculated by dividing agreements between coders by agreements plus disagreements and multiplying by 100. The obtained IRR on the initial inclusion criteria was 97%. Any disagreements between the coders were discussed until they came to consensus or a third coder reviewed the disagreements and made the final decision.

Application of the basic design standards. Following initial inclusion/exclusion procedures, each article was reviewed to determine whether or not it met the basic design standards developed by the What Works Clearinghouse (WWC; Kratochwill et al., 2010), adapted by Maggin, Briesch, and Chafouleas (2013). From here forward, the procedures described only apply to the 32 studies from the original meta-analysis that included adolescent- and adult-aged participants. Because studies that did not meet the basic standards were excluded by Ninci et al. (2015) prior to selection for the current meta-analysis, none of the 32 studies reviewed for the current meta-analysis did not meet the basic design standards.

Inter-rater reliability on design standards. Two independent coders reviewed the 32 articles (100%), as a component of the larger meta-analysis (Ninci et al., 2015), to determine whether each of the studies met the design standards, met them with reservations, or did not meet the design standards. Averaged across the six design standards, the IOA was 88%.

Variable Coding
For the purposes of this meta-analysis, the 32 studies that either met the design standards or met them with reservations were summarized based on the following variables: independent variable, participant diagnosis, and dependent variable. If a study variable did not fit in the classified categories or did not provide descriptions regarding these variables, the study variable was coded as “OTHERS.” For each moderator, anything categorized as “OTHERS” was excluded from further analysis due to the heterogeneity of that category.

The independent variable was divided into four levels, determined by the following criteria: (a) if a study used a type of video modeling, the intervention variable was coded as “VM”; (b) if a study used a type of audio cueing defined as that an instructor provided verbal instructions to the participant, the intervention variable was coded as “AC”; (c) if a study used a type of behavioral in-vivo instruction, including reinforcement, prompting, and prompt fading, the intervention variable was coded as “BIV”; and (d) if a study used a type of visual cue, including schedules, pictorial task analysis, and social stories, the independent variable was coded as “VC.” Studies that used video modeling, audio cueing, or visual cues often tended to use a type of behavioral in-vivo techniques. Therefore, if a study used video modeling, audio cueing, or visual cues in combination with behavioral in-vivo techniques, the intervention variable was coded as the former (i.e., video modeling, audio cueing, visual cues).

The participant’s diagnosis was coded for the following variables: (a) if a study reported
that the participant had autism spectrum disorder or autism, the participant’s diagnosis variable was coded as “AU”; (b) if a study reported that the participant had high-functioning autism or Asperger Syndrome, the participant’s diagnosis variable was coded as “HFAAS”; and (c) if a study reported that the participant had autism with a comorbid diagnosis of intellectual disability (i.e., the participant’s IQ was below 70) or if the participant’s adaptive behavior scores were two or more years delayed compared to the age equivalent score, the participant’s diagnosis variable was coded as “AUDD.”

The dependent variable included the following four levels: (a) if a study targeted self-help skills, such as toileting, cooking, hygiene, bathing, tooth-brushing, dressing, and independent eating, the outcome variable was coded as “SH”; (b) if a study measured house chore related skills, including cleaning and laundry, the outcome variable was coded as “HC”; (c) if a study targeted any employment skills, the outcome variable was coded as “ES”; and (d) if a study measured community access skills, such as transportation use, banking, and shopping, the outcome variable was coded as “COMMACC.”

Inter-rater reliability on variable coding. Variables were coded as a component of the prior meta-analysis (Ninci et al., 2015); however, to increase the proportion coded by two raters, additional articles were coded for the current review. Two independent coders reviewed a total of 16 studies (50%) to establish IRR for the four variable categories. The obtained IRR was 94%, ranging from 80 to 100% by variable.

Data Extraction and Analysis

To calculate effect sizes, data were extracted from each graph in each study included in this review. In the event that data measured behaviors of the participants who did not have ASD, or behaviors other than functional living skills, these data were excluded from the analysis. Graphs that did not meet the basic design standards were also excluded from further data analysis. Two adjacent phases were contrasted at a time (e.g., A1 vs. B1, A2 vs. B2; Parker et al., 2010b). A rank-order technique was used to extract data from each graph (Parker, Vannest, & Davis, 2011a). A data point plotted at the lowest point across the adjacent phases was ranked number 1. A data point plotted at the second lowest point across the adjacent phases was ranked number 2. This iterative process was continued until a data point plotted at the highest point across the phases was ranked. In the event that two or more data points were tied or plotted at the same point, the same rank number was assigned to those data points. If a study targeted a behavior intended to decrease, data were ranked in reverse order.

Inter-rater reliability for data extraction. Data were extracted as a component of the prior meta-analysis (Ninci et al., 2015). To increase the proportion coded by two raters, additional data were extracted for the current review. The second independent coder extracted a total of 869 data points from each graph across 13 studies (40%). IRR was calculated using a point-to-point percentage of agreement to determine correspondence of point-to-point data points across phases. The obtained IRR was 88.2%, ranging from 50 to 100% by experimental design.

Effect size calculation. A non-parametric Tau-U effect size (Parker et al., 2011b) was used to calculate effect sizes. In this review, Tau software developed through the Maple platform was used to calculate an individual effect size for each phase contrast, as well as an omnibus effect size (Davis & Davis, 2014). Tau-U scores ranged from $-1.0$ to $1.0$. If Tau-U scores were larger than 0.0, it indicated improvement between the two phases. If Tau-U scores were smaller than 0.0, it indicated that there was a deteriorating data set between the phases. Tau-U scores can be interpreted as following: (1) if Tau-U score ranges from 0 to .62, it indicates a small effect; (2) if Tau-U score ranges from .63 to .92, it indicates a moderate effect; and (3) if Tau-U score ranges from .93 to 1.00, it indicates a large effect (Parker et al., 2011a).

After the Tau calculation, the Kruskal-Wallis one-way analysis of variance was calculated as an omnibus test to determine if there were any differences in the group distributions (Kruskal & Wallis, 1952). The Kruskal-Wallis is a nonparametric analysis that is typically used to evaluate group differences when data do not meet standard ANOVA assumptions of normality (Elliott, & Hynan, 2011), as is often
the case with single-case data. If a statistically significant difference was found for any of the variables, a Dunn post-hoc test was planned to evaluate the pair-wise combinations (Dunn, 1964).

Additional analysis on video modeling. In order to determine the circumstances and participants for whom functional living skills are most effective, we conducted analyses to determine effect sizes for four independent variables, three participant diagnoses, and four dependent variables; however, we also conducted an additional analysis so we could draw more fine grained conclusions regarding for whom and for what specific dependent variables each intervention is most effective. Because half of the studies (N = 16) evaluated in this review used video modeling as an interventional technique, an additional analysis was conducted to identify how the effects of video modeling were differentiated by participant diagnosis and the dependent variable. Additional analyses were not conducted on other independent variables identified in this meta-analysis (i.e., audio cueing, behavioral in-vivo instruction, and visual cues) due to the small number of studies that have used these interventions to teach functional living skills to adolescent- and adult-aged participants with ASD.

### Results

Data from this study yielded 162 separate AB contrasts from 32 unique studies with 86 participants. Within all experiments analyzed, a wide range of Tau-U effect sizes were identified (i.e., .372 to 1.000), while a majority of studies resulted in moderate effects or strong effects (i.e., .724–1.000). In addition to the overall effect sizes across the experiments, analyses for variables were conducted.

#### Independent Variable

Analysis of the independent variable had four unique variables (see Table 1). In order to generate an effect size for each category, the following contrasts were analyzed: in the audio cueing (AC) category, 18 contrasts across 6 studies were evaluated; in the video modeling (VM) category, 83 contrasts across 16 studies were evaluated; in the behavioral in-vivo (BIV) category, 38 contrasts across 7 studies were evaluated; and in the visual cues (VC) category, 10 contrasts across 3 studies were evaluated. Within this analysis, Tau-U effect sizes ranged from a moderate effect of .890 CI[.838, 942] for video modeling to a strong effect of 1.000 CI[.826, 1.000] for visual cues. The Krukal-Wallis showed no sta-

### TABLE 1

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Number of Studies</th>
<th>Number of Study Participants</th>
<th>Number of Contrasts</th>
<th>Group Tau [CI95]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio cueing</td>
<td>6</td>
<td>11</td>
<td>18</td>
<td>.940 [.828, 1.000]</td>
</tr>
<tr>
<td>Video modeling</td>
<td>16</td>
<td>46</td>
<td>83</td>
<td>.890 [.838, 942]</td>
</tr>
<tr>
<td>Behavioral in-vivo</td>
<td>7</td>
<td>23</td>
<td>38</td>
<td>.917 [.851, .984]</td>
</tr>
<tr>
<td>Visual cues</td>
<td>3</td>
<td>8</td>
<td>10</td>
<td>1.000 [.826, 1.000]</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autism and either mental retardation</td>
<td>25</td>
<td>60</td>
<td>120</td>
<td>.922 [.879, .965]</td>
</tr>
<tr>
<td>or intellectual disability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASD or autism</td>
<td>9</td>
<td>17</td>
<td>23</td>
<td>.918 [.818, 1.000]</td>
</tr>
<tr>
<td>Cognitively high-functioning autism or Asperger syndrome</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>.724 [.586, 1.000]</td>
</tr>
<tr>
<td>Dependent variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment skills</td>
<td>18</td>
<td>50</td>
<td>82</td>
<td>.906 [.855, .957]</td>
</tr>
<tr>
<td>Self-help skills</td>
<td>6</td>
<td>15</td>
<td>34</td>
<td>1.000 [.905, 1.000]</td>
</tr>
<tr>
<td>House chores</td>
<td>5</td>
<td>16</td>
<td>26</td>
<td>.851 [.757, .945]</td>
</tr>
<tr>
<td>Community access skills</td>
<td>5</td>
<td>10</td>
<td>19</td>
<td>.929 [.841, 1.000]</td>
</tr>
</tbody>
</table>
statistically significant differences between studies based on the independent variable categorization \((p = .044)\).

**Participant Diagnosis**

Three unique variables were categorized within the diagnosis category (see Table 1). The following contrasts were analyzed in order to generate an effect size for each category: in the autism and intellectual disability (AUIDD) category, 120 contrasts across 25 studies were evaluated; in the ASD or autism (AU) category, 23 contrasts across 9 studies were evaluated; and in the cognitively high-functioning autism or Asperger syndrome (HFAAS) category, 9 contrasts across 3 studies were evaluated. Within this analysis, Tau-U effect sizes ranged from a moderate effect of 0.724 CI\(_{95}\) [.586, 1.000] for a diagnosis categorized as cognitively high-functioning autism or Asperger syndrome to a strong effect of 0.922 CI\(_{95}\) [.879, .965] for a diagnosis categorized as autism and intellectual disability, indicating moderate effects were found for each of the three of the participant diagnosis categories. The Kruskal-Wallis analysis indicated no statistically significant differences between studies based on diagnostic categorization \((p = .749)\).

**Dependent Variable**

Within the dependent variable category, four unique variables were categorized (see Table 1). The following contrasts were analyzed in order to generate an effect size for each category: in the employment skills (ES) category, 82 contrasts across 18 studies were evaluated; in the self-help skills (SH) category, 34 contrasts across 6 studies were evaluated; in the house chores (HC) category, 26 contrasts across 5 studies were evaluated; and in the community access skills (COMMACC) category, 19 contrasts across 5 studies were evaluated. Within
this analysis, Tau-U effect sizes ranged from a moderate effect of .851 CI95 [0.757, .945] for a dependent variable categorized as house chores to a strong effect of 1.000 CI95 [.905, 1.000] for a dependent variable categorized as self-help skills, indicating that strong effects were found for each of the four dependent variable categories. The Kruskal-Wallis indicated statistically significant differences between studies based on the dependent variable categorization (p < 0.01).

**Additional analysis on video modeling: Dependent variable.** For the additional analysis using video modeling only, four unique variables were categorized within the diagnosis category (see Table 2). In order to generate an effect size for each category, the following contrasts were analyzed: in the employment skills (ES) category, a total of 40 contrasts across 7 studies were evaluated; in the self-help skills (SH) category, a total of 13 contrasts across 4 studies were evaluated; in the house chores (HC) category, a total of 26 contrasts across 5 studies were evaluated; and in the community access skills (COMMACC) category, 3 contrasts across 2 studies were evaluated. As a result of these analyses, Tau-U effect sizes ranged from a moderate effect of 0.720 CI95 [.355, 1.000] for the community access skills category, to a strong effect of 1.000 CI95 [.852, 1.000] for the self-help skills category. The Kruskal-Wallis analysis indicated statistically significant differences between dependent variables (p < 0.01).

**Table 3**

<table>
<thead>
<tr>
<th>Group Comparisons</th>
<th>Difference in Average Ranks</th>
<th>Cutoff at Alpha = 0.05</th>
<th>Significance Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community access skills-Employment skills</td>
<td>33.3542</td>
<td>37.6111</td>
<td>**</td>
</tr>
<tr>
<td>Community access skills-Self-help skills</td>
<td>44.1667</td>
<td>40.2439</td>
<td></td>
</tr>
<tr>
<td>Community access skills-House chores</td>
<td>28.0513</td>
<td>38.3111</td>
<td></td>
</tr>
<tr>
<td>Employment skills- House chores</td>
<td>5.3029</td>
<td>15.8281</td>
<td></td>
</tr>
<tr>
<td>Employment skills- Self-help skills</td>
<td>10.8125</td>
<td>20.0590</td>
<td></td>
</tr>
<tr>
<td>House chores- Self-help skills</td>
<td>16.1154</td>
<td>21.3426</td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

The purpose of this meta-analysis was to investigate the magnitudes of effect of educational interventions for teaching functional living skills to adolescent-and adult-aged individuals with ASD. Specifically, we investigated the following moderators: (a) participant diagnoses, (b) independent variables, and (c) dependent variables. In addition, we also conducted an additional analysis with studies that used video modeling as an intervention in order to identify how the effects of video modeling were differentiated by participant diagnosis and the dependent variable. Additional analyses were not conducted with audio cueing, behavioral in-vivo instruction, or visual cues due to the paucity of research that has been conducted using these interventions to teach functional living skills to adults and adolescents with ASD. Lastly, we investigated whether there were any statistically significant differences between levels or categories of these moderators. Although researchers have recently begun to focus on the best way to treat adolescent-and adult-aged individuals with ASD (Cimera & Cowan, 2009), it is still unclear how to best foster the independent living skills in this population. Ninci and colleagues (2015) recently conducted a meta-analysis on functional living skills in individuals with ASD; however, because they did not focus on adolescent-and adult-aged individuals, questions still remained. This meta-analysis contributes to the literature.
base by answering questions about which educational interventions are most effective for teaching functional living skills to adolescent- and adult-aged individuals with ASD.

The first variable that was investigated was the independent variable, or the type of educational intervention used to teach functional living skills. Specifically, this study explored the following four independent variables: audio cueing, video modeling, behavioral in-vivo, and visual cues. Strong effects were found for all four independent variables, indicating that they are all effective for teaching functional living skills to adolescent- and adult-age individuals with ASD. However, because three of the interventions (i.e., audio cueing, visual cues, and behavioral in-vivo interventions) included seven or fewer studies, these results should be viewed with caution. Video modeling included the most studies, which suggests that there is a strong research base to support its use with adolescents and adults with ASD to teach functional living skills.

The second variable that was evaluated was participant diagnoses. While Tau-U analyses showed moderate to strong effects for all three diagnoses, the strongest effects were found for individuals with autism and intellectual disability and for individuals with ASD or autism. One possible reason that individuals with cognitively high-functioning autism or Asperger syndrome did not show as strong effects as those with the two aforementioned diagnoses could be the limited number of studies conducted with this population. This is not surprising given that adaptive behavior functioning has a strong negative correlation with autism symptomatology (Kenworthy, Case, Harms, Martin, & Wallace, 2010); in other words, individuals with high-functioning autism often have strong adaptive behavior, or functional living skills.

In order to gain further insight into the population for whom video modeling is most effective, we conducted an additional analysis using only studies that used video modeling as an intervention. The results of this analysis demonstrated moderate effects for all three participant diagnoses; however, the strongest effect was found for autism and intellectual disability. This indicates that video modeling is a viable intervention to teach functional living skills to adults and adolescents diagnosed with autism and a comorbid intellectual disability. Considering the moderately large number of studies (N = 13) and participants (N = 35) used in the analysis to calculate the effect size for individuals with autism and intellectual disability, readers can be fairly confident in the reliability of these results. However, the other two diagnoses, ASD or autism and cognitively high-functioning autism or Asperger syndrome, included a very limited number of studies and participants. As a result, the effects found for these two diagnoses should be interpreted with caution.

The third moderator that we investigated was the dependent variable in the study, or the type of functional living skill that was taught. This meta-analysis explored the following four dependent variables: employment skills, self-help skills, house chores, and community access skills. All four dependent variables showed strong effects, indicating that these are all skills that can be successfully taught to individuals with ASD. Out of the four dependent variables, employment skills was the skill most often investigated. This is expected, given that adolescent- and adult-age individuals were investigated in this meta-analysis. Considering that individuals with ASD are at a high risk for unemployment (Shattuck et al., 2012), it is encouraging that researchers are teaching employment skills and that the interventions are shown to be effective.

In order to gain a better understanding of the effects of specific interventions on the four functional living skills identified in this meta-analysis, we conducted an additional analysis on the dependent variables using only studies that used video modeling as the independent variable. The results of this analysis indicated that video modeling is moderately effective for teaching employment skills, house chores, and community access skills to adolescents and adults with ASD and strongly effective for teaching self-help skills to this population. Considering the ease with which video modeling can be implemented by practitioners (Carnahan, Basham, Christman, & Hollingshead, 2012), it is encouraging that the results demonstrated it is an effective intervention for teaching functional living skills to adults and adolescents with autism. It should be noted, however, that the community access skills category included a very small
number of studies and therefore, the results should be interpreted with caution.

As is the case with most meta-analytical research, this study does include some limitations. First, as was the case in previous work (Ninci et al., 2015), we did not analyze data on generalization or maintenance conditions. This leaves questions as to whether the effects found for each of the interventions, particularly video modeling, will continue over time. Another limitation is the inclusion of only published studies. Because published studies typically include positive results, meta-analyses consisting of only published studies may result in biased conclusions (Duval & Tweedie, 2000). The results of this meta-analysis, therefore, should be interpreted with caution. Finally, while statistical analyses are the preferred method for evaluating single-case research and contribute to the identification of evidenced-based practices, they do not provide information regarding the context in which the data were collected (Brossart et al., 2014). Thus, readers are encouraged to consider the context in which the data were collected for the individual studies included in this meta-analysis when interpreting the statistical analyses presented.

This study also identifies several areas that need additional research. It would be beneficial to analyze data on generalization and maintenance conditions in future meta-analyses in order to determine long-term effects. Visual cues and audio cueing should also be investigated in future research to determine if this intervention will continue to show strong effects when investigated with additional participants. Finally, while employment skills were investigated in a number of studies (more than all of the other functional living skills combined), there remains a paucity of research on self-help skills, house chores, and skills related to accessing the community. Future research should focus on interventions designed to increase these skills among adolescent- and adult-aged individuals with ASD.

In conclusion, this meta-analysis contributes to the literature base by investigating the magnitudes of effect of four educational interventions on teaching functional living skills to adolescents and adults with ASD. In addition, this meta-analysis also answered questions about the effects of using video modeling to teach functional living skills to this population. Despite some limitations, this study has provided evidence to demonstrate that functional living skills can be effectively taught to adolescents and adults with autism using educational interventions. This is encouraging as it provides evidence that the skill deficits often experienced by adolescents and adults with autism can be remediated, which may lead to more positive outcomes for this population.

References

Studies included in the analysis are listed here: https://u.tamu.edu/adoladultADLmetarefs


*Vineland Adaptive Behavior Scales* (2nd ed.). Circle 
Sun, X., & Allison, C. (2010). A review of the preva-
ence of autism spectrum disorder in Asia. *Re-
doi:10.1016/j.rasd.2009.10.003
Volkmar, F. R., Sparrow, S. S., Goudreau, D., Cic-
Social deficits in autism: An operational approach 
using the *Vineland Adaptive Behavior Scales*. *Journal 
of the American Academy of Child & Adolescent 
Psychiatry, 26*, 156–161. doi:10.1097/00004583-
199305000-00020
and vocational skills among individuals with 
autism spectrum disorder: Predictors, impact, 
and interventions. *Review Journal of Autism and 
Developmental Disorders, 1*, 266–275. doi:10.1007/
s40489-014-0024-7

Received: 30 June 2016
Initial Acceptance: 23 August 2016
Final Acceptance: 25 September 2016
Effects of Modeling, Story Templates, and Self-Graphing in the Use of Story Elements by Students with Moderate Intellectual Disability

Robert Pennington
University of Louisville

Mallory Koehler
Oldham County Schools

Abstract: There is limited research on teaching narrative writing skills to students with moderate to severe intellectual disability. In the current study, we used a multiple probe across participants single case design to evaluate the effects of an intervention package comprised of modeling, story templates, and self-graphing, on the inclusion of story elements in the written narrative of three middle school students with MSD. The data suggested that the intervention package was effective and resulted in partial maintenance of targeted skills. Implications for future research are discussed.

Teaching students to communicate effectively through writing presents one of the greatest instructional challenges for educational professionals. Data from the National Assessment of Educational Progress (2011) suggest that teachers have struggled to identify and apply effective strategies to develop writing skills to a majority of learners. In fact, the report indicated that over 70% of students in eighth and twelfth grades failed to meet proficiency in the area of written expression. Unfortunately, this bleak narrative worsens for young writers with disabilities as they often fall behind their peers in the development of written language (Mayes & Calhoun, 2007; Myles et al., 2003; Poplin, Gray, Larsen, Bankowski, & Mehring, 1980). This general lack of success in developing competent writers may have broad and deleterious effects, as struggling writers will likely face repeated and daunting writing requirements in higher education and employment settings.

Despite the importance of writing for all students, there has been little attention paid to the development of written expression by persons with moderate to severe intellectual disability (MSD; Joseph & Konrad, 2008; Pennington & Delano, 2014). Students with MSD may be at greater risk for failing to acquire written language skills due to differences in memory, learning rates, attention, and communication functioning (Heward, 2012). These differences may hinder students’ efforts in organization, vocabulary usage, and adherence to syntax and grammar conventions during writing. The resulting inability to generate cohesive written products may reduce students’ participation in general education settings, as teachers often expect the demonstration of knowledge through written expression. Furthermore, writing instruction may have a positive impact on communication in that the expansion of written responses may produce gains in vocal expression. For example, Pennington, Collins, Stenhoff, Turner, and Gunselman (2014) demonstrated that following instruction on how to construct simple story narratives, elementary-aged students with autism spectrum disorder (ASD) and MSD were able to vocally tell stories.

Fortunately, researchers have identified a collection of research-based practices for improving the writing of students from kindergarten to high school (Graham & Harris, 2016; Graham & Perin, 2007), including the provision of (a) frequent opportunities to write, (b) opportunities for students to increase comprehension by writing about text or information, (c) highly motivating routines for daily writ-
ing, (d) prewriting activities, (e) supports for students writing as they compose, (f) models and explicit instruction to teach writing skills and processes, (g) setting specific product goals, and (h) assistive technology supports. Several research teams have employed these practices in teaching writing to students with MSD. For example, Collins, Branson, Hall, and Rankin (2001) used explicit instruction to teach letters containing four elements: (a) heading, (b) greeting, (c) content body, and (d) closing. During instruction, a peer used the system of least prompts (SLP) procedure to teach participants to include each letter element. During SLP, the peer applied a prompting hierarchy in response to student errors or a failure to respond that included (a) an opportunity to respond independently, (b) a verbal prompt, (c) a gesture prompt, (d) a model prompt, and (e) physical guidance. All three participants generated a letter that included the targeted components but demonstrated only partial maintenance of skills at 2 and 4 weeks after instruction. Similarly, Pennington, Delano, and Scott (2014) employed explicit instruction and setting product goals in teaching three young adults, ages 18–21, with MSD to include targeted components within resume cover letters to potential employers. During instruction, the researchers used modeling, the SLP procedure, and taught students to evaluate and graph their daily performance. During instruction, the three participants learned to use all six components and maintained the use of components several weeks after the termination of intervention.

A few researchers have targeted story writing for students with MSD. Story writing is an important skill for young learners as it provides an opportunity to access attention from the listener or reader. It also provides a rich context for refining communication skills and writing conventions. In fact, the generation of a cohesive narrative has been targeted for students across grade levels by the Common Core State Standards (CCSS; National Governors Association & Council of Chief School Officers, 2010). These standards emphasize (a) writing for multiple purposes, (b) producing well-organized text, (c) writing to recall and extend knowledge about a topic, and (d) using short and extended writing across a range of audiences (Graham & Harris 2013). In a series of studies (i.e., Pennington, Ault, Schuster & Sanders, 2011; Pennington, Stenhoff, Gibson, & Ballou, 2012; Pennington, Collins et al., 2014), researchers taught nine children, ages 6–10, to construct simple stories using prompting and assistive technology. In these investigations, the researchers used predictable story patterns (e.g., [character] lived in the [locale], [character] [action], [character] was [emotion]) and simultaneous prompting (Gibson & Schuster, 1992) to teach participants to select whole words from an array to complete a story. Participants acquired the skills, demonstrated maintenance, and some responded across untrained topographies (i.e., speaking, handwriting). These studies suggest that students with MSD can learn to construct simple stories when given a model and the vocabulary from which to choose. Unfortunately, they provide little direction on teaching students to generate their own stories.

In the current investigation, we sought to build upon the previous work in teaching story writing to students with MSD by targeting the generation of stories in the absence of computer-based arrays (e.g., Pennington et al., 2014). The instruction of generative responses is critical in that it frees the writer from those responses pre-programmed into a word array and provides the opportunity for the writer to tell their own story. To date, there is a dearth of research on how to develop a generative repertoire for students with MSD.

In addition, we targeted the instruction of middle school students with MSD. Only one other research team has addressed writing instruction for middle school students with MSD. Mims, Lee, Browder, Zakas, and Flynn (2012) targeted writing as a part of a comprehensive literacy package and used a time delay procedure to teach students to complete sentence fill-ins to “write” an opinion statement about a passage. In the current investigation, we employed strategies consistent with recommendations for teaching writing to middle school students with MSD (Graham & Perin, 2007). First, we used prewriting and modeling strategies by presenting opportunities to practice a writing structure in the context for motivating video-based stories. Researchers have employed videos in teaching a wide range of
skills to students with MSD. In most cases, videos were used to model the accurate performance of participant responses (Bellini & Akullian, 2007); we used video clips to show students the story elements within a narrative and as stimuli about which to model story writing responses. Second, we used an explicit response prompting strategy (SLP) during instruction. The SLP procedure has been used to teach a wide range of skills to students with MSD (Doyle, Wolery, Ault, & Gast, 1988), including several in the area of written expression (Collins et al., 2001; Pennington, Delano, & Scott, 2014; Pennington, Flick, & Smith-Wehr, in press). Finally, we taught students to graph their progress towards product goals. Recently, two research teams included self-graphing as a part of an effective instructional package for teaching writing skills to student with MSD (Pennington, Delano, & Scott, 2014; Pennington, Saadati, Scott, & Welch, 2014). In light of this proposed extension of the current literature, we sought to answer the following research question: Is there a functional relation between an intervention package comprised of video models, story templates, and self graphing on the number of story elements included in the written narratives three middle school students with MSD?

Method

Participants

Three males, ages 12–13, with moderate intellectual disability participated in the study. All three students performed at the low or very low range of performance on standardized writing and language assessments. Ryan was a 12-year-old white male with moderate intellectual disability and ASD and in the seventh grade. His most recent evaluation indicated a full-scale IQ score of 41 (WISC-IV), a written expression score of 40 (KTEA), and a core language score of 71 (CELF-V). Duncan was a 13-year-old white male with moderate intellectual disability and in the sixth grade. His most recent evaluation indicated a full-scale IQ score of 52 (WISC-IV), a written expression score of 49 (KTEA), and a core language score of 72 (CELF-V). Duncan received free and reduced lunch. Their special education teacher recommended the students for participation in the study because of their deficits in the production of cohesive narratives. All three students had basic spelling skills and could produce simple written sentences about picture or a single event (e.g., I went to the movies).

The student’s special education teacher conducted all baseline and intervention sessions. She held a bachelor’s degree and teaching certification in Moderate to Severe Disability (MSD) and had 2 years of teaching experience in a middle school.

Settings and Materials

We conducted all sessions in a self-contained special education classroom for students with MSD that served students in sixth-eighth grade. Students moved in and out of the classroom as they spent part of their day in general education classrooms. Sessions were conducted in a one to one instructional arrangement, whereas the teacher sat across from the student at a small semi-circle shaped table. Other students received instruction from paraprofessional in other areas of the classroom. Each session lasted approximately 15 to 20 min.

Across all sessions, two students used pencil and paper to generate narratives. One student, Ryan, typed his responses on a computer tablet using a word processing software program (i.e., Spell Better). During intervention sessions, students were presented with five brief video clips selected from YouTube and displayed on a computer tablet. We selected each video vignette based the following criterion: (a) contained all targeted story elements, (b) lasted less than 4 min, and (c) contained what the researcher and teacher perceived as a potentially engaging ending (e.g., high level of action, humor). We also randomly presented two story templates to model simple story writing (see Figure 1).
Each template contained sentence starters followed by blank space (e.g., once there was______________). In addition, we used a histogram in which students plotted their daily progress by coloring the number of blocks that corresponded with the number of correct responses. Above the graph, we created and placed a checklist of the five story elements (see Figure 2). The graph and list were presented after the student had finished independent writing.

Data Collection

We collected data on responses during independent writing practice and on the occurrence of five story elements within each student’s written narrative. Story elements were scored as occurring if it was present and related to the rest of the narrative. Furthermore, each element must have been included within a phrase or sentence (i.e., minimum subject + verb) and errors in punctuation or capitalization were ignored. Story elements included (a) character, (b) setting, (c) a first event, (d) a related second event, and (e) an emotion. A character was defined as a person or animal that performed an activity in the story. A setting was defined as a statement of a place or time in which the story takes place. The first and second events were defined as something the character did or something that happened that affected the character. Finally, an emotion was scored as a statement of how the character felt. We assigned a single point for the inclusion of each element. Therefore, participant scores could range from 0 to 5 points. Data were reported and graphed as number of story elements included during writing probes.

Reliability. The first author, a university researcher, collected both dependent and independent reliability data. For dependent variable reliability data, he independently scored students products and compared his score

![Figure 1. Sample Story Template.](image)

![Figure 2. Self-graphing histogram and checklist.](image)
against that of the teacher. He then calculated interobserver agreement (IOA) by dividing the number of agreements by the sum of agreements and disagreements and then multiplying by 100. For Ryan, we collected IOA data for 75% of baseline sessions and 100% of intervention sessions. IOA was 90% across sessions. For Jay, we collected IOA for 50% of baseline and 100% of intervention sessions. IOA was 100% across sessions. For Duncan, we collected data for 42% of baseline and 100% of intervention sessions. IOA was 96% across sessions.

We also collected procedural integrity data across all baseline, intervention, and maintenance conditions. During sessions, the first author sat behind the teacher and student dyad and recorded the teacher’s adherence to the planned instructional steps. For Ryan, we collected data for 25% of baseline sessions, 60% of intervention sessions, and 40% of maintenance probes. For Jay, we collected data for 50% of baseline sessions, 50% of intervention sessions, and 66% of maintenance probes. For Duncan, we collected data for 57% of baseline sessions, 50% of intervention sessions. We calculated procedural integrity to be 100% across all sessions.

**Experimental Design and Procedures**

We used a concurrent multiple probe (MP) across participants (Horner & Baer, 1978) to evaluate the effects of a writing intervention package on the participants’ use of story components in their writing samples. We used the MP design for its capacity to demonstrate experimental control while reducing the need for students to repeatedly engage in difficult writing tasks without intervention. First, we collected baseline data across all three participants. Subsequently, we introduced the writing package to the first participant and continued until he met a predetermined criterion (i.e., 5 elements for 3 consecutive sessions). This criterion was selected based on a previous study involving generative writing for students with MSD in which students maintained responses following intervention (Pennington, Delano, & Scott, 2014). We then introduced the intervention to the next participant and continued until he met criterion. Finally, we introduced intervention to the final participant and again continued until criterion was met. Following the termination of each participant’s intervention condition, we conducted probes to assess maintenance.

**Teacher Training**

Prior to baseline and intervention conditions, the researcher trained the special education teacher to implement procedures. Prior to training, the researcher sent a sequenced list of teacher behaviors via email for review. At the onset of a single day of training, the researcher reviewed the steps and modeled implementation of the steps. Finally, the teacher was asked to implement the steps with the researcher until she demonstrated performance with 100% accuracy. The researcher sat across from the teacher and scored her performance using a researcher develop checklist. Furthermore, the teacher was permitted to reference the written list of steps during instruction.

**Baseline Condition**

During baseline sessions, the teacher presented writing materials and obtained the participant’s attention. Subsequently, she delivered the directive to write a story and directed the participant to vocally indicate when he/she was finished writing. She also informed the participant that they could ask for help in spelling. If the participant discontinued writing for 3 min, the teacher asked the student if he/she is finished. If the student indicated “no” they were permitted to continue until they indicated completion or another 3-minute pause elapsed at which time the teacher termination the session. Once the student indicated completion or the session was terminated, the teacher delivered praise for on-task behavior.

**Writing Intervention Package**

Each intervention session involved three instructional segments including (a) a story model and checklist review, (b) independent student practice, and (c) feedback plus self-graphing. The teacher conducted sessions 3 to 4 days a week.

**Story video model and checklist review.** At the beginning of each intervention session, the
teacher stated, “It’s time to work on story writing” and displayed a brief video depicting a cartoon short story. After watching the video, the teacher presented a template and stated, “A story has several different parts.” The teacher then directed the student to the template and stated, “A story needs a character and a character is who the story is about. Who was the main character in the story?” If the participant answered the question correctly within 5 s, the teacher provided praise, wrote his response on the story template and read the section aloud. If the student did not respond, indicated that he did not know, or answered incorrectly, the teacher initiated the following prompt sequence. First, the teacher replayed a segment of the video that displayed the targeted elements and then repeated the question. If the student answered correctly, the teacher provided praise, wrote his response on the story template and read the response aloud. If the student did not provide the correct response, the teacher pointed to the video, verbally provided the answer, wrote his response on the story template, and read the response aloud. The teacher repeated the prompting and feedback steps across all of the targeted elements. Table 1 depicts the definitions and questions provided during the instruction of each targeted element. Upon completion of the template the teacher stated, “We wrote a story together” and then read the entire story aloud. She then stated, “Remember your story has several important parts. Let’s review them.” Subsequently, she pointed to each component on the checklist (Figure 2), stated the name of the components and asked the student to identify the corresponding component from the story (e.g., character-Jack-Jack). If the student did not respond within 5 s, or made an error, the teacher provided the correct response.

**Independent student practice.** After reading the co-constructed story, the teacher removed the template and presented the directive to write a different story. She reminded the participant to include all of the important story elements, request assistance for spelling when needed, and indicate when he was finished. If the student stopped writing for 3 min, the teacher asked him if he was finished. If he indicated that he was not finished, he was permitted to proceed or continue without writing for an additional 3 minutes. The practice session was terminated following 6 consecutive minutes without student writing.

**Feedback plus self-graphing.** Upon completion of the students’ practice narrative, the teacher presented a checklist, graph, and the statement “OK, let’s see of you have all of the parts.” The teacher then asked the student to identify the occurrence or nonoccurrence of each component. If the student accurately identified the occurrence or nonoccurrence of an element, the teacher provided praise. If the student made an error, the teacher provided the correct answer. The teacher and student then tallied the number of elements included and the student plotted the number of elements on his graph.

**Maintenance probes.** Following completion of the intervention, the teacher conducted
maintenance probes using procedures identical to those used during baseline sessions.

**Results**

Results are displayed in Figure 3.

*Ryan.* Prior to intervention, Ryan’s performance data were variable and indicated an average of 2.25 elements per session. On session 2, his data show the use of four elements, which reflects his generation of a narrative...
about a recent trick or treating trip. During intervention, Ryan Jay demonstrated a steadily increasing trend and met criterion within five sessions. Immediately following intervention, during maintenance sessions 10 and 11, Ryan included three and five elements respectively. During maintenance session 12, which occurred 3 weeks after intervention, he included a single element. Finally, Ryan used all five elements at 2.5 and 3 months following intervention.

Jay. Prior to intervention, Jay’s performance data were variable and included an average of one element per session. On session 2, he included four elements in a story about a trip to Dollywood. During intervention, Jay demonstrated a steadily increasing trend and met criterion within 6 sessions. During maintenance sessions, Jay included three elements one day following the termination intervention. He included five elements at 5 weeks following intervention and at 9 weeks following intervention, he responded with a request “to go on a cruise during spring break” resulting in the score of 0.

Duncan. Prior to intervention, Duncan included an average of .42 elements per session. Following the introduction of intervention, Duncan’s data reflect an immediate change in responding. He met criterion within eight sessions. At 4 weeks following intervention Duncan used 4 elements during a maintenance probe.

Non-parametric Measure of Effect

We used Tau-U (Parker, Vannest, Davis, & Sauber, 2010) to measure non-overlap between phases and to provide a non-parametric measure of effect. Tau-U improves upon previous indices (e.g., percent of non-overlapping data points [PND], percent of data exceeding a median trend [PEM-T]) in its integration of overlap and trend (Parker, Vannest, & Davis, 2014). We first measured non-overlap between baseline and intervention conditions for all three participants. We then calculated an overall effect across all participants. For Ryan, Jay, and Duncan we calculated the effect to be .55, .53 and .98, respectively. The weighted average across all participants was .70. These findings reflect small effects (i.e., < .65) of Ryan and Jay, a strong effect for Duncan (> .93–1.00), and an overall medium effect across the three participants (.66–.92).

Discussion

In the current investigation, we addressed the research question: Is there a functional relation between an intervention package comprised of video models, story templates, and self graphing, and the number of story elements included in narratives by three middle school students with MSD? Our findings suggest a functional relation as the package was effective for all three participants in increasing the use of targeted story elements during intervention. Furthermore, the participants demonstrated variable but greater than baseline levels of responding during maintenance probes. These findings add to the limited but emerging body of literature suggesting that students with MSD can benefit from systematic writing interventions and are of little surprise as systematic and explicit instructional strategies have been deemed effective for students without MSD (e.g., Self-Regulated Strategy Development, Graham & Harris, 1989; sentence combining, Saddler, Behforooz, & Asaro, 2008).

In the current investigation, we organized instruction around predictable writing routines. During instruction, students completed sentences within story templates that served as models of the target written product. Sentence completion has been previously been identified as a critical task within the scope and sequence of instruction in written expression (Kame’enui & Simmons, 1990). These tasks have three important instructional features. First, they reduce the complexity of writing tasks, as they require students to have a limited understanding of syntax. Second, they provide intra stimulus models of correct sentence structure. Finally, the sentence stems or frames within these tasks may serve to mediate generalization (Stokes & Baer, 1978) as they can be applied in the generation of multiple sentences. Interestingly, we observed that the students included some of the sentence stems within their own narratives.

In addition, we presented video examples of story narratives to teach students to identify critical story elements. In previous work with students with MSD, Pennington and col-
leagues (2011, 2012, 2014) modeled correct student responses (i.e., selection of words) but did not provide models of complete story narratives prior to instruction. It is plausible that students may have acquired story-writing responses without actually understanding what constitutes a story narrative. This may have contributed to the limited variation within student responses across the studies. In the current study, two of the participants, Ryan and Jay, demonstrated wide variation in their stories. Each of their stories differed from the video exemplar and stories written on previous days. Duncan stories adhered closely to the narrative in the video models. Furthermore, our use of video during writing instruction for this population was somewhat novel as only one other research team has used videos during writing instruction for this population (Kinney, Vedora, & Stromer, 2003).

Finally, it is important to note that Ryan’s use of the particular writing software (i.e., Spell Better) presented challenges during the study. This software included word prediction technology that presented a choice of words above the keyboard display. We observed on multiple occasions that Ryan seemed to “tire” of typing out each word and sporadically selected multiple words from the three choice array of predicted words. Often these brief shifts to selection based responding resulting in syntactically incorrect phrases. Researchers and practitioners might consider the inhibitive effects of predictive software for students without the skills to use it effectively.

Limitations

Despite our generally positive outcomes, we must acknowledge limitations to the current analysis. First, as mentioned above, Ryan’s sporadic use of word prediction software questions limited our ability to fully assess his generative performance. Since the feature was available across baseline and intervention conditions, our data suggest the intervention likely contributed to changes in responding but that these effects should be considered with caution. Future research should isolate these effects by avoiding the use of predicative software or ensuring that students can use it discriminately.

Second, we only targeted the inclusion of story elements within their narratives. Though story elements are critical in learning to produce cohesive narratives, they only reflect one set of responses within the complex repertoire required to be a competent writer. In the current investigation, students made frequent errors in spelling and grammar that may have reduced the readability by natural readers (e.g., peers, teachers, parents). Despite this limitation, we purport that these data are a critical early step toward developing more comprehensive writing packages.

Finally, our data reflect the performance of only three participants. We caution practitioners in their interpretation of our results as written expression is a complex construct and its development is highly responsive to individual learning histories. That being said, our findings are consistent with an explicit and systematic approach that has been demonstrated to be effective across a range of individuals (Graham & Harris, 2016).

Conclusion

The increased expectation for all students to make progress in the areas academic areas including written expression has served as a catalyst for changes in the educational programming for students with MSD. Unfortunately, teaching these students to write is new territory for researchers and practitioners. The current study adds to the emerging body of literature in this novel area and may reveal to future researchers potential next steps towards the development of more comprehensive programming. We purport that the rationale for the incorporation of writing into the curriculum of persons with MSD goes beyond requirement to meet national or state standards and towards the goal self-determination. At its core, written expression is another form of communication in which individuals use to demonstrate their knowledge, interact with others through electronic media, and express their feelings.

References

spectrum disorders. Exceptional Children, 73, 264–287.


Abstract: The purpose of the present study was to compare the effectiveness and efficiency of simultaneous prompting with and without video modeling in teaching food preparation skills to four participants with autism spectrum disorder, whose ages ranged from 5 to 6 years old. An adapted alternating treatment single-case experimental design was used to compare the procedures used in the study. The study findings demonstrated that both procedures were equally effective in promoting the acquisition of the food preparation skills for three of the participants with autism. Simultaneous prompting with video modeling was more effective for the remaining participant. Mixed results were obtained regarding the effectiveness parameters. Although mixed results were obtained during the study regarding effectiveness, it can be said that there was only a small difference in favor of simultaneous prompting using video modeling.

The number of children with autism spectrum disorder (ASD) is gradually increasing around the world. According to the U.S. Centers for Disease Control and Prevention (CDC, 2015), approximately one child in every 68 in the United States is diagnosed with ASD. Prevalence of ASD in Asia, Europe and North America average about 1% (CDC, 2015). ASD is a neurodevelopmental disorder that impacts many aspects of social and communicative functioning, with associated impact on cognitive and adaptive functioning (Klin, 2011). In addition to deficits in social and communicative functioning, restricted, repetitive, stereotyped patterns of behavior negatively affect the learning of adaptive behaviors in children with ASD (American Psychiatric Association, 2013; Bodfish, 2011). Adaptive functioning describes the degree to which individuals are able to function independently in their daily lives. The acquisition of independent living skills, such as self-care (e.g. dressing), adaptive skills (e.g. food preparation) should be a priority for children with ASD (Klin et al., 2007; Taylor, DeQuinzio, & Stine, 2014). Independence in daily life improves quality of life, not only for the individual with ASD, but also for his or her family (Taylor et al., 2014).

Some effective instructional procedures have been developed to teach adaptive skills to children with ASD. Nowadays, a variety of interventions and procedures have been developed to meet the requirements associated with the deficits of these children. Using an effective procedure, and selecting an efficient and socially valid approach, is important for the education of children with ASD (Kurt & Tekin-Iftar, 2008; Schuster, Griffen, & Wolery, 1992; Tekin-Iftar & Kircaali-Iftar, 2006).

Increasing the number of children diagnosed, the variety of services and seeking of effective instructions are required to identify and describe the evidence-based practices for children with ASD (Odom et al., 2003). The National Professional Development Center (NPDC, 2015) and National Autism Center (NAC, 2015), which are organizations dedicated to developing evidence-based practices, report various practices for teaching children with ASD. According to their reports, discrete
trial teaching and video modeling are two of the evidence-based practices for teaching various skills to children with ASD.

Simultaneous prompting (SP), also referred to as errorless learning, is a response-prompting procedure that has been successfully implemented within discrete trial training sessions. SP is a procedure designed to minimize errors. Consequently, SP trials are presented at a zero-second delay; that is, the implementer provides the prompt immediately after providing the direction (Gibson & Schuster, 1992). Since the implementer does not provide an opportunity for the participant to complete the task independently during the intervention sessions, separate probe sessions are conducted to examine the acquisition of the skill. Therefore, the implementer can make a decision regarding whether or not skill acquisition has occurred. During probe sessions, after the attention of the participant is secured, the teacher can present a direction or discriminative stimulus and wait 3–5 seconds for a response. Correct responses may result in praise and incorrect responses, may result in error correction, or the instructor may ignore them altogether (Gibson & Schuster, 1992). Probe sessions can either be conducted prior to every intervention session, or spread throughout the week (Morse & Schuster, 2004).

Studies have shown that SP is an effective instructional procedure for teaching discrete skills, such as object naming (MacFarland-Smith, Schuster, & Stevens, 1993), identifying community signs (Singleton, Schuster, & Ault, 1995), identifying first-aid materials (Tekin-Iftar, Acar, & Kurt, 2003), identifying occupations (Dogan & Tekin-Iftar, 2002), naming features on a map (Gursel, Tekin-Iftar, & Bozkurt, 2006), and identifying transportation vehicles (Reichow & Wolery, 2009). SP is also effective for teaching chained skills, such as opening a keyed lock on a locker (Fetko, Schuster, Harley, & Collins, 1999), assembling shipping boxes (Maciag, Schuster, Collins, & Cooper, 2000), playing with a baby doll (Colozzi, Ward, & Crotty, 2008), and hand washing (Parrott, Schuster, Collins, & Gassaway, 2000) from preschool age to adults diagnosed with ASD and intellectual disabilities, as well as with typically developing children. Furthermore, professionals, nonprofessionals, teachers, family members, and peers have been shown how to implement SP correctly (Gibson & Schuster, 1992; Leaf, Sheldon, & Sherman, 2010; Morse & Schuster, 2004; Parker & Schuster, 2002; Tekin & Kircali-Iftar, 2002). However, further research is still needed to increase the instructional efficiency of SP, in terms of reducing the number of trials to meet the criteria, the number of errors, and total training time (Akmanoglu, Kurt, & Kapan, 2015; Kurt & Tekin-Iftar, 2008; Morse & Schuster, 2004; Reichow & Wolery, 2009; Waugh, Alberto, & Fredrick, 2011).

Video modeling (VM) is an evidence-based practice that has been developed, based on Bandura’s social learning theory (as cited in Bellini & Akullian, 2007). VM incorporates making a video recording of adults, peers, or him or herself completing a behavior or task correctly. The child watches video of the behavior and then imitates the behavior independently (Bellini & Akullian, 2007; Coy & Hermansen, 2007; Nikopoulos & Kenan, 2006). Research has shown that VM is an effective instructional strategy for teaching social-communication, play, and independent life skills in children and adolescents with ASD and developmental disability (Bellini & Akullian, 2007; Mason, Ganz, Parker, Burke, & Camargo, 2012). Although VM is a very effective intervention for the acquisition of these target behaviors, some studies have reported mixed results (Apple, Billingsley, & Schwartz, 2005; D’Ateno, Mangiapanello, & Taylor, 2003; Taylor, Levin, & Jasper, 1999). Researchers describe mixed results with the sole use of VM, which on certain occasions may not offer positive gains for all participants (Delano, 2007; Mason et al., 2012). Therefore, in order to obtain a higher level of achievement VM may need to be combined with other interventions or components.

VM is an intervention that is frequently combined with other interventions or components (Bellini & Akullian, 2007; Genc-Tosun & Kurt, 2014). In the literature, there are many studies that investigate the combined effects of VM with activity schedules (Blum-Dimaya, Reeve, Reeve, & Hoch, 2010), social stories (Sansosti & Powell-Smith, 2008), prompting (Akmanoglu & Tekin-Iftar, 2011; Graves, Collins, & Schuster, 2005), and reinforcement (Charlop-Christy & Daneshvar, 2003). In addition to effectiveness studies,
there are some studies that have compared the instructional efficiency of different types of interventions while using with VM. For example, two studies have demonstrated that implementing VM before PECS and least-to-most prompting strategies, have resulted in the faster acquisition of behavior. Future studies need to examine the effects of VM on instructional efficiency of other intervention procedures (Cihak, Smith, Cornett, & Coleman, 2012; Murzynski & Bourret, 2007).

Although VM and SP are often effective when implemented individually, it may be that participants could learn skills more rapidly and with less instructor time when the two interventions are combined. Therefore, in this study, we investigated the effects of VM on the instructional efficiency of SP. The purpose of the present study was to compare the effectiveness and efficiency on the acquisition of chained adaptive skills in children with ASD and of implementing SP alone versus implementing VM and SP together. The following research questions were asked: Which intervention(s) are more effective in teaching adaptive skills to children with autism spectrum disorder? Which intervention(s) are more efficient in reaching the criterion, regarding (a) the number of training sessions, (b) the number of training trials, (c) the percentage of errors, and (d) total training time? Furthermore, the social validity of this study was investigated via a questionnaire to the parents and teachers.

Method

Participants

Four male preschool-aged children diagnosed with ASD participated in the study. All participants attended individual or group training programs at a developmental support unit of a research institute in Turkey. All the participants had been diagnosed with ASD by child psychiatrist at a university hospital. Additional confirmatory support for diagnosis was obtained by having the teachers complete the Turkish version of the Gilliam Autism Rating Scale (GARS-2-TR), adapted and standardized by Diken, Ardic, and Diken (2011). Adaptive scores were not available for these participants; however, our observations determined that all of the participants had age-appropriate fine and gross motor skills, that they exhibited motor imitation, and that they did not demonstrate any challenging behaviors. None of the participants had a history of systematic instruction with VM but each had experience with SP. Prerequisite skills for the participants were as follows: (a) follow one step verbal instructions, (b) have fine and gross motor skills sufficient for learning target food-preparation skills (described below), (c) able to concentrate on watching television for 2 minutes, (d) able to attend to visual and aural stimuli for 5 minutes.

Ersin was a 5-year old boy with ASD who was diagnosed at a university hospital when he was 2.5 years old and his GARS score was 93. According to his teacher report and the authors’ observations, he did not use any spontaneous language, did not imitate verbalizations, and did not make any vocalizations. However, he did follow one or two-step verbal instructions and attend to an activity for 10 minutes. He could perform basic self-help skills, such as eating, washing, dressing and toileting, independently. On weekdays, he attended group training at the unit for three hours a day. Cem was 6-year old and had been diagnosed with ASD at a university hospital when he was 3 years old and his GARS score was 90. According to his teacher report and the authors’ observations, he could independently use two or three-word sentences, respond to simple questions (e.g., “What is this?” “Who is she?”), follow verbal instructions, and attend to an activity for 15 minutes. He had difficulty with social and communication skills (e.g., he had limited eye contact, he had difficulty responding to greetings and interaction with peers). He could perform basic self-help skills, such as eating, washing, dressing and toileting, independently. He attended individual training sessions at the unit for two hours per week. Burak was a 5-year old boy with ASD. He was diagnosed with ASD at a university hospital when he was 3 years old and his GARS score was 93. According to the teacher report and the authors’ observations, he could imitate one- to two-word sentences, respond to simple questions (e.g., “How are you?” “What is your name?”), follow one or two-step verbal instructions, and attend to an activity for 15 minutes. He had difficulty with social and communica-
tion skills (e.g., he had limited eye contact; he had difficulty interacting with his peers). He was able to perform basic self-help skills such as eating, dressing, toileting independently. He attended group training at the unit on weekdays for three hours each day. Melih was a 6-year-old boy, diagnosed with ASD at the age of 2.5 and his GARS score was 91. Observation indicated he did not speak, imitate others, or make any vocalizations. He was able to follow simple instructions (e.g., “Give me red ball!”) and attend to an activity for 10 minutes. He had difficulty initiating social interactions and used some gestures and signs. He could perform basic self-help skills such as eating, washing, dressing, and toileting with help verbal prompting. He attended group training at the unit on weekdays for three hours each day.

**Trainer and observer.** The researcher (i.e., first author) conducted all experimental sessions. She is a Master’s student in special education and has a minimum of 2 years experience with systematic instruction. The reliability observer was also a graduate student in a Master’s program in special education.

**Setting and Materials**

All probe, intervention and maintenance sessions were conducted in the developmentally support unit of the research institute in a one-to-one teaching arrangement. All sessions were video-recorded. SP treatment was conducted in the unit’s cafeteria. The cafeteria had three cabinets, one refrigerator, two long tables, and twenty chairs. The VM sessions were conducted in one of the one-on-one study rooms (4m × 3m). The study rooms had a television, filing cabinets and a DVD player. The participant and researcher sat on the same side of the table to watch the video. The same settings were used both for the video clips and training sessions. Generalization sessions were undertaken by their parents at home in the kitchen. The same materials were used during all sessions.

Materials used during the study included a video model, a computer, a television with a 72-cm screen, a video CD player, the objects used to teach target skills (e.g., milk, cocoa, glass, bread, cheese, hot chocolate, popcorn), a video camera, and data collection forms. The same setting were used both video clips and training sessions.

**Modeling videos.** Peer video modeling was used in this study. Peer modeling is a type of VM which target behaviors are performed by a familiar peer, such as siblings, classmates or unknown individuals (Coy & Hermansen, 2007). The first author created video clips, which depicted the scenario from the third-person perspective. In other words, the viewpoint is similar to that of an observer. A typically developing peer modeled the appropriate responses for one of the target skills, until the participant accomplished the task correctly. The peer attended second grade at an elementary school and was unknown to the participants. The length of each clip was between 30 and 70 seconds. Six professionals who had experience with special education and VM procedures watched the clips to ensure that the model fully and comprehensively demonstrated the target skills.

**Experimental Design**

An adapted alternating-treatments design was used to investigate the relationship between the dependent variables and independent variables and the efficiency of VM+SP procedure, and SP alone, on teaching target skills for participants with ASD (Holcombe, Wolery, & Gast, 1994). Using an adapted alternating-treatments design, each of two similar but functionally different dependent variables, receive a different independent variable (Holcombe et al., 1994). The dependent variable of the study was the percentage of correct responses on the steps of the task analyses of the target skills. Two similarly complex food-preparation tasks were implemented with each participant for each independent variable. The independent variables of the study were VM+SP and SP alone. The implementation of VM+SP and SP alone with the target skills was counterbalanced across four participants. Rapid alternation of the interventions was administered by allowing at least one hour between the sessions. For example, SP alone was implemented with one target skill, and at least one hour later VM+SP was implemented with other target skill to the same participant.
Selection of Target Skills

The researchers initially selected adaptive skills (preparing a food or drink) for each participant by asking the participants’ teachers what food-preparation sequences were needed (and not mastered) for each participant. The food-preparation sequences were similar but functionally different and consisted of the same number of steps and similar levels of complexity. The pool of skills was divided into three sets. Each set contained two chained skills. Later, parents choose two target skills for each of their children from the pool. The lack of prior mastery of each food-preparation sequence was confirmed via the collection of baseline data, as described below.

Dependent Measures

Two dependent measures were evaluated in this study. The first was the percentage of correct responses during baseline and intermittent probe sessions. Correct responses were defined as correctly performing a step of the task analyses within 4 seconds, and incorrect responses were defined as incorrectly performing a step of the task analyses, not completing it in 4 seconds, or performing a step differently from the task analyses. The lack of prior mastery of each food-preparation sequence was confirmed via the collection of baseline data, as described below.

Data Collection Probe Sessions (Baseline and Intermittent Probe Sessions)

There were baseline and intermittent probe sessions in the study. Baseline sessions were conducted prior to teaching chained skills and continued until stable data and a minimum of three data points were obtained for each food-preparation sequence and for each participant. A baseline session was performed as follows: The implementer delivered a specific attention cue to secure the participants’ attention (e.g., “Melih would you like to eat some cereal?”), and after receiving an affirmative response from the participant through either eye contact or a gesture, the implementer praised the participant (e.g., “Good!”). The implementer than delivered the task direction (e.g., “Let’s prepare some cereal for you.”), waited 4 seconds for the participant’s response and delivered verbal praise for every step of the task analysis (e.g., “Good job!”). The trial ended when the participant made an error and all remaining steps were scored as incorrect. Intermittent probe sessions were conducted every other day immediately before implementing training sessions in each food-preparation sequence in order to test acquisition.

Training Sessions

A total task format was used to teach the chained skills. Physical prompting was used for each participant and the implementer stood behind or next to the participant. Correct responses resulted in verbal reinforcement. Participants’ cooperation skills were reinforced verbally at the end of each session. The instructional procedures, SP with/without VM, were randomly assigned to the chained skills. Training was provided until a 100% correct response was obtained on at least three consecutive intermittent probe sessions.

**TABLE 1**

Task Analysis for Chocolate Milk

<table>
<thead>
<tr>
<th>Steps in the Task Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Open milk bottle</td>
</tr>
<tr>
<td>2. Hold milk bottle</td>
</tr>
<tr>
<td>3. Pour milk in glass</td>
</tr>
<tr>
<td>4. Put milk bottle on table</td>
</tr>
<tr>
<td>5. Close milk bottle</td>
</tr>
<tr>
<td>6. Open chocolate powder box</td>
</tr>
<tr>
<td>7. Put the cover of chocolate box on table</td>
</tr>
<tr>
<td>8. Get table spoon</td>
</tr>
<tr>
<td>9. Fill table spoon with chocolate powder</td>
</tr>
<tr>
<td>10. Pour chocolate in glass</td>
</tr>
<tr>
<td>11. Stir the milk with spoon</td>
</tr>
<tr>
<td>12. Put table spoon on tray</td>
</tr>
<tr>
<td>13. Get the cover of chocolate powder</td>
</tr>
<tr>
<td>14. Close chocolate box</td>
</tr>
</tbody>
</table>
Simultaneous prompting alone. After the baseline sessions, the implementer started to teach chained skill, as SP sessions. The instructional trial with SP was delivered as follows: the implementer secured the participant’s attention (e.g., “Ersin, would you like to make a chocolate milk?”). After receiving an affirmative response to the question, the implementer verbally reinforced the participant (e.g., “Good!”) and delivered the task direction (e.g., “Ersin, make a chocolate milk please!”). The implementer immediately delivered full physical prompt to the participant. The implementer held the participant’s hand and completed the task analyses together. So, the participant did not have a chance to make any incorrect responses. Correct responses by the participant were reinforced verbally (e.g., “Perfect!”). At the end of the session, as natural reinforcement, the implementer was given a hazelnut-spread sandwich to eat.

Video modeling with simultaneous prompting. After consistent data had been obtained during baseline sessions, the implementer initiated instruction with VM+SP. The intervention sessions were conducted as follows: The implementer and the participant watched the video together, If the participant did not watch, the implementer delivered verbally direction (e.g., “Ersin, please watch the television!”), after watching the video; the participant was given verbal reinforcements (e.g., “You are watching television very well!”). The implementer and the participant went to the cafeteria together, then the implementer secured the participant’s attention (e.g., “Ersin, would you like to make a sandwich?”). After receiving an affirmative response to the question, the implementer verbally reinforced the participant (e.g., “Good!”) and delivered the task direction (e.g., “Ersin please make a sandwich.”). The implementer immediately delivered a physical prompt and kept providing the prompt to the participant until the task completed. A correct response by the participant was reinforced verbally (e.g., “Good job, Ersin!”). At the end of the session the implementer said the participant could eat the food.

Maintenance Sessions

Maintenance probe sessions conducted 1, 2 and 4 weeks following the criterion were met by all participants, apart from Ersin. The third maintenance session for Ersin was conducted 90 days after first maintenance session, because of his family’s summer vacation. Maintenance sessions were conducted just like the probe session, apart from the reinforcement schedule. The verbal reinforcements were delivered on solely the basis of the correct completion of the tasks.

Generalization Sessions

Generalization sessions were conducted in the participants’ homes with their parents. Generalization sessions were performed in a pre-test and post-test format. A pre-test generalization session was carried out immediately after baseline sessions and post-test generalization sessions were carried out directly after last intermittent probe session.

Post-test generalization sessions conducted once every 2 weeks. Generalization probe trials were conducted in the same manner as the maintenance trials and the same response definitions were used.

Reliability

Dependent variable reliability and procedural reliability data were collected during at least 30% of the sessions. A graduate student in special education served as an independent observer and collected the reliability data. Inter-observer agreement was calculated by dividing the total number of agreements between the experimenter and the observer by the total number of agreements plus disagreements and multiplying by 100% (Tawney & Gast, 1984; Tekin-Iftar & Kircaali-Iftar, 2006). The mean percentage of agreement was 99% (range = 94%–100%) for all participants. Inter-observer agreement data for all participants during all sessions and conditions are shown in Table 2.

Procedural reliability data were collected for 30% of training sessions to determine whether the independent variable was being used as initially planned. Procedural reliability was calculated, by dividing the number of steps accurately completed, by the total number of training steps and multiplying by 100 (Billingsley, White, & Munson, 1980). The procedural reliability was average 98% (range = 80%–100%) across four
participants. Procedural reliability data for all participants are shown in Table 3.

Results

Effectiveness Data

Figures 1 through 4 show the percentages of correct responses during baseline, intervention, generalization, and maintenance sessions for all participants across instructional procedures. The findings of this study indicated that both independent variables were effective. As can be seen from the figures, both SP alone and VM+SP were equally effective on promoting acquisition of the food-preparation sequence for three participants.

TABLE 2

Interobserver Agreement Data for All Participants during All Sessions and Conditions

<table>
<thead>
<tr>
<th>Participants</th>
<th>Target Skills</th>
<th>Baseline</th>
<th>Intermittent</th>
<th>Probe</th>
<th>Intervention</th>
<th>Generalization</th>
<th>Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ersin</td>
<td>Preparing chocolate milk</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Making a cheese sandwich</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Cem</td>
<td>Making popcorn</td>
<td>94%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preparing hot chocolate</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Burak</td>
<td>Making a fruit drink</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Making a hazelnut-spread sandwich</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Melih</td>
<td>Making a hazelnut-spread sandwich</td>
<td>100%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preparing cereal</td>
<td>100%</td>
<td>95%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Note. R = range.

TABLE 3

Procedural Reliability Data for All Participants during All Sessions and Conditions

<table>
<thead>
<tr>
<th>Participants</th>
<th>Target Skills</th>
<th>Baseline</th>
<th>Intermittent</th>
<th>Probe</th>
<th>Intervention</th>
<th>Generalization</th>
<th>Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ersin</td>
<td>Preparing chocolate milk</td>
<td>100%</td>
<td>98%</td>
<td>99%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Making a cheese sandwich</td>
<td>100%</td>
<td>100%</td>
<td>99%</td>
<td>90%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Cem</td>
<td>Making popcorn</td>
<td>100%</td>
<td>99%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preparing hot chocolate</td>
<td>100%</td>
<td>97%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Burak</td>
<td>Making a fruit drink</td>
<td>100%</td>
<td>97%</td>
<td>98%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Making a hazelnut-spread sandwich</td>
<td>100%</td>
<td>100%</td>
<td>99%</td>
<td>90%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Melih</td>
<td>Making a hazelnut-spread sandwich</td>
<td>100%</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preparing cereal</td>
<td>100%</td>
<td>99%</td>
<td>99%</td>
<td>80%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Note. R = range.
However, these results were not replicated with the fourth participant, Melih.

Ersin did not perform any steps of either food-preparation sequence during baseline. When intervention sessions started with both procedures, the trend and level of each chained skill immediately improved. Ersin mastered the preparation of chocolate milk with SP alone over 16 trials and making sandwich VM+SP over 14 trials.

Cem provided between 0% and 10% correct responses for making popcorn and did not perform any correct responses for preparing hot chocolate during the baseline. When in-

![Figure 1. Percent of correct responses during baseline, intermittent probe, generalization maintenance sessions for Ersin.](Image)

![Figure 2. Percent of correct responses during baseline, intermittent probe, generalization maintenance sessions for Cem.](Image)
Intervention sessions were initiated with both procedures, the trend and level of each chained skill immediately improved. During intervention, he acquired 100% accuracy for preparing hot chocolate with VM+SP over 12 trials and making popcorn with SP alone over 14 trials.

Burak did not provide correct responses for either chained skill during the baseline. When intervention sessions were initiated with both procedures, the trend and level of chained skill immediately improved. During intervention, he mastered the preparation of a fruit drink with SP alone over 16 trials and preparing a hazelnut-spread sandwich with VM+SP over 22 trials.

Melih provided between 0% and 25% correct responses for cereal making skills with VM+SP alone, and between 0% and 8% correct responses for preparing hazelnut-spread sand-
which with SP alone, during the baseline. During intervention, he acquired 100% accuracy for preparing cereal with VM+SP over 12 trials but he did not meet the criteria for making a hazelnut-spread sandwich taught by SP alone. Unfortunately, the study was discontinued at that point due to summer vacation.

Another finding of the present study was that the participants’ maintained the chained skills, with the exception of the sequence Melih had not mastered, after the intervention phase had ended. Furthermore, participants’ high levels of performance were also observed in the generalization sessions, which they performed at home with their parents.

Efficiency Data

Three primary variables were measured: the number of trials to reach criterion, number of incorrect responses to reach criterion, and the total amount of time taken to reach the criterion. The efficiency data shows that, SP with VM was more efficient than SP alone with regard to the number of training sessions and trials to reach the criterion for Ersin, Cem and Melih. However, SP alone was more efficient than VM+SP with regard to the number of training sessions and trials to reach the criterion for Burak. On the other hand, for two participants (Burak and Ersin), SP alone seemed to be more efficient in terms of the percentage of incorrect responses. For Cem, no differences were found between the two dependent variables, in terms of the percentage of incorrect responses. The SP alone procedure seemed to be more efficient than VM+SP for all participants, in terms of the total amount time taken to reach the criterion. Efficiency results for each participant are presented in Table 4.

Social Validation

The parents (n = 6) and teachers (n = 3) of the participants individually completed a social validity questionnaire at the end of the intervention to share their opinions regarding the purpose of the study, the appropriateness of the procedures, and the significance of the changes observed in the chained skills. The researchers developed a different Social Validity Form for teachers and parents, which included yes/no and open-ended questions. The teachers and parents found the intervention to be effective and acceptable. The parents reported that they would let their children prepare food that they learned as a result of this study. In addition, the parents said that participating in the generalization sessions of the study and implementing them with their children made them happy. The teachers reported that they had experience with SP procedure, but they had no experience of teaching with VM. Consequently, some of them stated that they would prefer to use both SP and VM separately. On the other hand, others reported if they had proper material and environment, they would use VM both alone and with SP in their class.

Discussion

The purpose of this study was to compare the effectiveness and efficiency of SP with VM and
SP alone, in the teaching of adaptive skills to children with ASD. The outcomes of this study revealed that both SP with VM and SP alone were effective in the teaching of adaptive skills to children with ASD. The effectiveness findings from this study are consistent with those of other studies, which separately examined the effectiveness of SP and VM in teaching chained skills (Akmanoglu & Batu, 2004; Alcantara, 1994; Bidwell & Rehfeldt, 2004; Colozzi et al., 2008; Dogan & Tekin-Iftar, 2002; Gibson & Schuster, 1992). Although the effectiveness data had positive findings, some significant points need to be discussed.

In this study, the criteria for both chained skills were met by three of the four participants. However, only Melih met 100% of the criteria in the preparation of cereal with VM and SP intervention procedures. Although Melih learned to make a hazelnut-spread sandwich with SP alone with 60% accuracy, the study had to be terminated because of summer vacation. When an adapted alternated treatments design is used to compare two independent variables in a study, using the more effective independent variable as a final phase is recommended (Holcombe et al., 1994). Therefore, to teach the making of a hazelnut-spread sandwich to Melih, by using more effective teaching strategy, VM and SP in this study, instructional sessions were planned. However, Melih left the institute and instruction could not be completed. The following points could explain his failure to learn this skill. First, the difficulty level of target behaviors may not have been equal for Melih, even though both target behaviors had the same number of steps in the task analysis. Second, during intervention sessions, the implementer observed that Melih was inclined to put the hazelnut-spread in his mouth directly, rather than spreading it on the bread first. As a result of this, he became stuck on the spreading step of the sandwich making, occasioning the failure of his performance.

To date, there are no published studies that have compared the effectiveness and efficiency of SP with VM and SP alone. The efficiency data presented above did not result in a strong conclusion as to which procedure was more efficient. That is, there was no apparent difference between SP with VM and SP alone, in terms of the intervention procedure of the efficiency variables. The results indicate that for three participants, SP with VM was more efficient than SP alone with regard to the number of training sessions and attempts to reach the criterion; for one participant, SP alone was more efficient than SP with VM in terms of all the efficiency variables. On the other hand, analysis of the percentage of errors to reach the criterion show that SP alone was more efficient than SP with VM for two participants and SP with VM was more efficient than SP alone for two participants. Meanwhile, SP alone, required less training time in all participants. Although mixed results on the efficiency measures were obtained in the study, it might be said that efficiency results showed only a small difference in favor of the VM + SP procedure. This is in general agreement with previous studies involving the combination of prompting and VM. Murzanski and Bourret (2007) stated that VM / least-to-most prompting was more efficient than least-to-most prompting alone, in teaching adaptive skills in terms of the number of trials and number of prompted steps. Cihak et al. (2012) found out that the VM + PECS intervention procedure required fewer training sessions than PECS procedure alone, for acquisition of independent communicative initiations.

The reason for the mixed results related to efficiency could be explained in this manner. First, SP might be an effective and efficient procedure without combining VM or other interventions. Secondly, the participants' individual characteristics may have affected the findings. For example, Melih might prefer to eat hazelnut-spread with a spoon, preventing him from completing the analysis task properly. Third, task difficulty of the selected chained skills might not establish an equivalent. Therefore, similar studies for teaching different skills (e.g., play skills, self-care skills) in different settings, by different implementers, could be conducted in the future.

The social validity findings of the study showed that both the participants' teachers and parents responded positively. The data was in agreement with previous studies that investigated the effects of SP in terms of social validity (Akmanoglu & Batu, 2004; Colozzi et al., 2008; Kurt & Tekin-Iftar, 2008; Waugh, Fredrick & Alberto, 2009). Conversely, limited studies have reported social validity outcomes
of VM (Bellini & Akullian, 2007); therefore, these findings enhance the existing literature regarding the measurement of the social validity of VM.

Generalization across two conditions (participants’ home and with their parents) was assessed in this study. Generalization is an essential challenge for individuals with ASD. Research indicates that VM facilitate generalization. However, only limited studies have conducted generalization sessions (Bellini & Akullian, 2007). Participants have demonstrated generalization across setting and people. Furthermore, most previous studies have not provided a measure of treatment fidelity for VM. This study was conducted with average of 98% (range: 80%–100%) across all sessions; therefore, these findings contribute to the current literature regarding the generalization outcomes and treatment fidelity.

Although this study contributes to the existing literature in some ways, there were a number of limitations should be mentioned. First of all, a small sample was used in this study. Therefore, to obtain consistent findings and generalizability of results, future researchers should include a larger sample size with different disabilities. Secondly, this study was conducted with a structured setting in the cafeteria of the unit, except for the generalization sessions that were implemented at participants’ homes. In other words, the sessions used in the study were not conducted in natural settings and occasions. Thirdly, to determine the performance levels of the participants during baseline and probe sessions, a single-opportunity method was used. Therefore, the probe data may not be accurate indicators of the participants’ actual performances. Future research may be improved by including multiple opportunities and an extended baseline phase.

Furthermore, future research should be conducted by adding different parameters to increase instructional efficiency. Additional studies should be conducted to investigate whether VM causes an increase on instructional efficiency of other response prompt teaching strategies (e.g., constant time delay, graduated guidance). There is no difference between SP with VM and SP alone intervention procedure, in terms of the effectiveness and efficiency variables. For this reason choosing best practice for their participants, according to participant and behavior features, is recommended to both teachers and implementers.

References


Charlop-Christy, M. H., & Daneshvar, S. (2005). Using video modeling to teach perspective taking...


Odom, S. L., Brown, W. H., Frey, T., Karasu, N.,


Received: 11 April 2016
Initial Acceptance: 8 June 2016
Final Acceptance: 1 September 2016
The Role of the Replacement Behavior in Function-Based Intervention

Linda M. Reeves
University of South Alabama

Jolenea B. Ferro
University of South Florida

John Umbreit
University of Arizona

Carl J. Liaupsin
University of Arizona

Abstract: Three students with autism spectrum disorder (ASD) who displayed off-task behavior participated in a two-phase study. In Phase 1, a functional behavioral assessment (FBA) was conducted for each student. In addition, an assessment of each student’s ability to perform the replacement behavior identified that none of the participants was able to do so. In Phase 2, two function-based interventions were designed for each student. Both interventions included identical reinforcement and extinction procedures but different antecedent conditions. One intervention prompted performance of the replacement behavior. The other directly taught the student how to perform the replacement behavior. Both interventions were tested during classroom activities. For these students, the intervention that included teaching the replacement behavior produced high levels of on-task behavior. In contrast, the intervention without direct instruction produced much lower on-task levels that were only marginally higher than baseline. Implications for intervention design are included.
Interventions include three types of intervention components: (a) antecedent conditions that set the occasion for a replacement behavior, (b) reinforcement procedures that support use of the replacement behavior, and (c) extinction procedures to be used when the target behavior occurs. (cf. Crone & Horner, 2003; Umbreit, Ferro, Liaupsin, & Lane, 2007).

The student’s ability to perform the replacement behavior independently (i.e., fluently enough to be reinforced naturally) plays a key role in intervention design. Interventionists need to determine whether a student’s failure to perform a replacement behavior reflects a performance deficit or an acquisition deficit. Students with a performance deficit fail to perform the replacement behavior at acceptable levels even though they have the knowledge and ability to perform the skill. In contrast, students with an acquisition deficit lack the knowledge and/or ability to perform the replacement behavior, even under optimal conditions (Gresham, Van, & Cook, 2006). When students have acquisition deficits, the intervention must include strategies to teach the student the replacement behavior (Umbreit et al., 2007).

The purpose of this study was to examine whether different outcomes would occur when the intervention for students with acquisition deficits did or did not include direct instruction in performance of the replacement behavior. The study was conducted in two phases. In Phase 1, researchers conducted a descriptive FBA for each student and verified that each had an acquisition deficit with respect to the replacement behavior. In Phase 2, two nearly identical interventions were designed for each student. One intervention prompted performance of the replacement behavior. The other directly taught the student how to perform the replacement behavior. Both interventions were tested during classroom activities.

**General Method**

**Participants and Setting**

The study was conducted in two elementary and one middle school classrooms in an urban public school district in the southwest. To be selected, participants needed to (a) be between the ages of 4 to 13, (b) be diagnosed with ASD, (c) have an IEP, (d) display chronic challenging behavior, (e) be placed in a classroom environment that demonstrated high levels of effective practices, and (f) fail to perform the replacement behavior independently. Three participants were selected.

**Participant 1.** Calli, age 12, was a sixth grader diagnosed with ASD and Speech/Language Impairment. She was placed in a self-contained class for students with Autism, but attended resource classes for science and language arts. According to her special education teacher, Calli read at grade level and received special education services in writing, math, speech and language, and counseling. She was referred for intervention by her special education teacher because her challenging behaviors were increasing in both frequency and intensity during language arts, and the language arts resource teacher was questioning the appropriateness of her continued placement in class.

The study took place during Calli’s language arts class. The class included 17 students, the teacher, and an educational assistant who accompanied Calli to provide academic and behavioral support. Students were arranged in five rows of five seats. The class routine included entry activities, daily assignment review, usually independent writing activities, and taking turns reading a story or participating in a class discussion.

**Participant 2.** Davis was a 5-year old student diagnosed with ASD. His kindergarten class included 28 students, his teacher, and a teacher’s assistant. According to his teacher, Davis knew all the letters in the alphabet, could count to 20, and write numbers up to 10. He also could identify most high frequency words for kindergarten and write his first and last name and simple sentences. Davis received special education services in speech and language. He was referred for intervention by his kindergarten teacher. She was concerned with Davis’ struggle to adapt to the class routine and, more specifically, his failure to follow directions and complete his work during math and language arts.

At the teacher’s request, the study took place during math. Students sat in groups of four to five at long tables. The routine consisted of completing a daily math activity that
usually involved counting objects and writing numbers, and often required the students to share materials or take turns. The students rotated to different math centers each day.

Participant 3. Gabe was a 6-year old first grader diagnosed with ASD. His class included 23 students and his teacher. Gabe received special education services in reading, written expression, and speech and language. His special education teacher referred him for intervention because of behaviors that included destroying his work, leaving class without permission, and threatening to leave school. Earlier in the school year, Gabe had been moved from another first grade class due to his challenging behaviors.

The study took place during language arts because Gabe’s teacher identified that subject and period as his most difficult time of the day. Students sat in small groups of four desks facing each other. The language arts class primarily consisted of writing activities that included engaging in bell work (e.g., writing the date, correcting errors, and making simple journal entries), independently completing worksheets, or writing about a topic while the teacher rotated among small groups for reading instruction.

Classroom Environment Assessment

Each student’s classroom environment was assessed to minimize the likelihood that challenging behaviors would be the result of ineffective classroom practices. For this purpose, the Classroom Environment Checklist (Ferro, Umbreit, & Liaupsin, 2008; available on request from the first author) was used. This instrument includes 33 items that address three domains: (a) classroom structure (12 items), (b) operating procedures and routines (13 items), and (c) classroom rules (8 items). Based on direct observation, each item is rated as in place, partially in place, or not in place. For the purposes of this study, credit was given only for items describing practices that were fully in place.

Each class was observed for at least 30 min by the first author and by a second independent observer. Each observer identified the number of items that were fully in place in each domain. The overall percentage of effective practices observed was calculated by dividing the total number of indicators observed by the total number of indicators in each domain, multiplying the results by 100%, and then averaging the scores in each domain to determine an overall score. IOA was calculated by dividing the number of agreements by the total number of indicators and multiplying the result by 100%.

Calli’s classroom environment was rated an overall score of 100% effective practices in place with 100% IOA. Davis’ classroom environment was rated an overall score of 100% effective practices in place with 97% IOA. Gabe’s classroom environment was rated an overall score of 93% effective practices in place with 94% IOA.

Behavioral Definitions

All three students exhibited behaviors that were considered off-task, although the topography of each student’s behavior varied. The replacement behavior for all participants was on-task behavior. For Calli, off-task behaviors included engaging in activities other than the assigned task, making rude comments or speaking rudely to teachers and peers (e.g., saying “I hate you”), complaining, raising her voice, screaming, whining or crying, banging on her desk, falling to the floor, and hitting her head. On-task behavior was defined as completing class assignments quietly, raising a hand to ask a question related to the topic or to ask for help, saying “no thank-you” when asked to participate, and “excuse me” to get someone’s attention.

For Davis, off-task behaviors included engaging in activities other than the assigned task, taking items from peers, crying, screaming, demanding items, and interrupting the teacher. On-task behavior was defined as completing the steps necessary to engage in activities independently, requesting items from peers, and getting his teacher’s attention by approaching and waiting to be acknowledged.

For Gabe, off-task behaviors included engaging in activities other than the assigned task, off-topic talking with peers, or leaving his seat to walk around the room for more than 10 s. On-task behavior was defined as engaging in assignments independently, following
directions, and raising his hand to ask for help, if needed.

**Procedure**

The study was conducted in two phases. In Phase 1, a descriptive FBA was conducted for each student. Each FBA included teacher interview, direct observations, and identification of the function(s) of each student’s target behavior. In addition, each student was assessed to determine whether he or she could independently perform the replacement behavior. In Phase 2, two interventions were designed for each student. Both interventions included identical reinforcement and extinction procedures. However, one intervention prompted performance of the replacement behavior, whereas the other directly taught the student how to perform the replacement behavior. Both interventions were then tested during class activities.

**Phase 1: Functional Behavioral Assessment and Assessment of Replacement Behavior**

**Data Collection and Analysis**

A FBA was conducted for each student to identify the antecedent conditions that set the occasion for their off-task behaviors, and the consequences that maintained these behaviors. Data were collected through a structured interview and direct observations in the classroom.

Structured staff interviews were completed with each teacher. Interviews were conducted using the *Preliminary Functional Assessment Survey* (Dunlap et al., 1993; available on request from the first author), a 22-item survey designed to gather information about the student’s behaviors of concern and the environmental conditions that contributed to their occurrence.

Structured observations involved the collection of antecedent-behavior-consequence (ABC) data (Bijou, Peterson, & Ault, 1968). These data were collected individually for each participant. Observations were conducted in the student’s classroom during naturally occurring activities in which the target behavior was most likely to occur. A-B-C data were collected until there was a clear pattern of antecedents and consequences related to the target behavior.

The function(s) of each student’s off-task behavior was identified by analyzing the FBA data using the *Function Matrix* (Umbreit et al., 2007). This visual tool assists users in determining whether the student is gaining access to something or escaping/avoiding something, or both. The user then identifies more specifically whether the student is gaining or escaping attention, tangibles/activities, or sensory consequences.

Each student’s ability to perform the replacement behavior independently prior to intervention was assessed using a two-step process. First, the steps involved in completing the replacement behavior were task analyzed (see Table 1 for a sample task analysis). Data were then collected on the number of steps each student performed independently. Second, a reinforcement contingency was added to see whether students would perform more steps under highly motivating conditions. The percentage of steps each participant completed independently was determined by dividing the number of steps performed inde-
Results

Results of the interviews, direct observations, function identification, and assessments of the replacement behavior are presented for each student.

**Calli.** A structured interview was conducted with Calli’s language arts teacher and the teaching assistant who accompanied her to class. The teacher was most concerned about Calli’s rude comments, which included Calli telling the teacher she was terrible and that she hated her. She reported that Calli acted out when work demands were placed on her. When off-task behaviors occurred, the teacher told her to stop, or offered her a break.

The assistant was most concerned about Calli’s disrespectful behavior and work avoidance. She indicated these behaviors occurred daily and had been escalating for the past month in language arts class. She believed Calli’s off-task behavior occurred when she was asked to complete an assignment she did not want to do. Disrespectful behavior occurred when she was given verbal cues to stop or the opportunity to take a break. If she continued, she eventually would be taken back to class or given an office referral.

A-B-C data were collected during five sessions for a total of 3 and 1/2 hours. Calli engaged in off-task behavior when (a) the teacher gave her instructions, (b) she was asked to share information, and (c) she was called on to participate. When these antecedents occurred, Calli engaged in other (non-directed) activities, complained, made rude comments to the teacher or assistant, screamed, banged her hand or pencil on her desk, hit herself on the head, fell to the floor, or left class without permission. In every case, Calli received attention from the teacher, assistant, or peers (e.g., additional explanations, reprimands, prompts to ask for help or a break, peers looking or laughing at her) and she also avoided her assignment or the request to contribute.

Data from the interviews and observations were organized into the Function Matrix to determine that Calli engaged in off-task behaviors when she was given an assignment and/or was called on to contribute. Off-task behavior enabled her to get attention from her teachers and peers and avoid the assignments and expectations.

Calli’s ability to perform the replacement behaviors independently was assessed over six days during her language arts class. Data were collected on whether Calli performed each step in the task analysis for the replacement behaviors: (a) task engagement, (b) raise hand to ask for help or an appropriate question, (c) say, “no thanks,” when asked to participate and, (d) “excuse me,” to get someone’s attention (see Table 1). Because task engagement included engaging in the assignment and sitting quietly, data were collected using 30 s whole intervals for at least 25 min. Data on the remaining social behaviors were collected using frequency data based on the number of opportunities during class.

For the first three days, the task analysis assessment showed Calli engaged in the replacement behavior an average of 14% (range 0–42%) of the time, and performed the steps of raising hand, saying “no thanks,” when called on, or “excuse me,” to get someone’s attention 0% of the time. For the next three days, a reinforcement contingency was added. Calli was told that, if she did her work and was respectful to her teachers and peers, she could leave class 15 min. early to either listen to her favorite music CD, or use the computer. Under the reinforcement contingency, Calli performed the steps in task engagement an average of 86% (range 67–98%) of the intervals, and said “no thanks,” 100% of the time based on one opportunity, but performed the steps to raise her hand 0% of opportunities, and got someone’s attention appropriately only 63% of opportunities. These data confirm a performance deficit with remaining engaged and saying “no thanks,” but an acquisition deficit with hand raising and getting someone’s attention.

**Davis.** Davis’ teacher reported that his academic abilities were at or above grade level, yet he often failed to complete nonpreferred activities. Sometimes he escalated to screaming, crying, and disrupting the class.
ported these behaviors occurred 5–10 times per day and were most likely to occur during centers. When Davis became disruptive, the teacher either sent him to the resource teacher’s class or to the principal’s office. The practice was unsuccessful. When he was sent to the resource teacher or principal, he would read a story or access the computer (preferred activities) and/or return to class with a picture of a solar system or map (a preferred item).

A-B-C data were collected during language arts and math centers for a total of three hours over four days. Off-task behavior occurred when Davis was asked to complete an activity during centers, and when he had to share materials or take turns with his peers. When these antecedents occurred, Davis interrupted the teacher or assistant to ask what he needed to do or to ask for a picture of a map or solar system. He interrupted the teacher by approaching them and saying “excuse me” repeatedly and progressively louder each time. The adults always gave him attention, reminded him that he needed to wait, and told him what to do. When another student had something Davis wanted, he either screamed that he needed the object, grabbed it from the student without asking, or interrupted the teacher to tell on the other student. The other students at his table either told him he needed to share, or argued with him until the teacher or assistant came over to get the students back on-task. Using the Function Matrix, the interview and observational data indicated Davis’ off-task behaviors during center activities enabled him to get teacher attention, assistance, and information on what he was supposed to do.

Davis’ ability to perform the replacement behaviors independently was assessed for four days during center activities. Data were collected on whether he completed each step in the task analysis of his replacement behavior (the full task analysis is available on request from the first author): (a) task engagement, (b) requesting items from peers appropriately, and (c) getting his teacher’s attention appropriately. For the first three days, data indicated that Davis independently performed the steps for engagement an average of 12% (range 0–20%) and the steps for both requesting items from peers and getting teacher’s attention 0% of the time under typical classroom conditions. On day four, a reinforcement contingency was arranged. Davis was told he could earn his choice of a picture of a solar system or a map. He continued to independently perform only 20% on the steps for engagement and 0% for requesting items from peers and getting teacher’s attention, clearly indicating an acquisition deficit across all three areas of the replacement behavior.

Gabe. Gabe’s teacher was most concerned about his failure to follow directions, complete his class work, and remain in class. She stated he was very capable of completing his assignments, but generally did not. She explained how on two occasions he had left class without permission and walked out of school.

A-B-C data were collected for 4 hours over four days. Off-task behavior occurred when Gabe was given an assignment to complete independently. Of the seven assignments he was expected to complete, Gabe completed three with additional teacher prompting and assistance, but did not even attempt the remaining assignments. Gabe’s teacher verbally prompted him 53 times by telling him or showing him where to write his name, reminding him to work or keep working, reading the question to him, or telling him where to write the answer. Gabe responded to 14 of those prompts (26%). Otherwise, he sat at his desk, played with small objects, interacted with peers inappropriately, or walked around the classroom. Using the Function Matrix, Gabe’s off-task behaviors occurred when he was asked to work independently and enabled him to either get additional prompting and assistance or to escape the assignment.

Ability to independently perform the steps in the replacement behavior was assessed four times during language arts. Data were collected on whether Gabe completed each step in his task analysis for on-task behavior (the full task analysis is available on request from the first author). For the first three days, Gabe performed an average of 14% of the steps (range 0–29%) under typically occurring classroom conditions. On day four, a reinforcement contingency was added. Gabe was told he could go to the library to read a book about volcanoes for the remainder of the period if he completed the class assignment. Under the reinforcement contingency, Gabe performed only 43% of the steps indepen-
ently, clearly indicating an acquisition defi-
cit.

Phase 2: Intervention Development and Testing

In Phase 2, two interventions were developed for each student. The two interventions included identical reinforcement and extinction procedures. However, one intervention prompted performance of the replacement behavior, whereas the other directly taught the student how to perform the replacement behavior. These interventions were then tested within the context of naturally occurring classroom activities and routines. For each student, the study took place over five weeks. It included 21 sessions for Calli, 25 sessions for Davis, and 22 sessions for Gabe.

Procedure

Each intervention included antecedent adjustments, reinforcement procedures, and extinction procedures. For each student, Intervention 1 involved prompting of the replacement behavior, but no direct instruction in how to perform it. Intervention 2 was identical but included direct instruction.

Calli. In Intervention 1, Calli was reminded of the behavioral expectations at the start of class. When the replacement behavior occurred, she was verbally praised. She also received a plus on her point sheet for every 5 min she was on-task. If Calli earned 80% of her points by a designated time, she got to leave class 10 min early for free time in her resource class, listening to music or playing on the computer. When the target behavior occurred, the teacher and assistant avoided responding. Intervention 2 was identical except that Calli was taught to raise her hand to get the teacher’s attention, and to get her assistant’s attention by looking in her vicinity and saying “excuse me.” To teach Calli the replacement behavior, the first author met with her for four 30-min sessions to provide instruction on examples and non-examples of how to ask for help, get someone’s attention, and greet others. She also developed (with Calli) a list of appropriate and inappropriate comments, and then taught Calli to keep the inappropriate comments to herself. A visual reminder of appropriate attention getting was also reviewed and placed in Calli’s vicinity.

Davis. For Intervention 1, Davis was reminded of the behavioral expectations at the start of the activity. When he was on-task, Davis was verbally praised and also earned a sticker for every 5 min he remained on task. When Davis earned five stickers during centers, he exchanged them for a daily token. When he had earned five daily tokens, he exchanged those for a picture of a solar system or map. If off-task behavior occurred, Davis was ignored and no preferred activities or items were offered as incentives to calm down. Intervention 2 was identical but included teaching Davis how to get his teachers attention appropriately and how to request items from peers. In addition, Davis was taught how to use visual reminders of the steps needed to complete his class activities. First, a folder containing pictures of the steps was created (e.g., write name, count, write, clean up.). Before each activity, the picture sequences were arranged in the order needed to complete the activity and then reviewed with Davis. As he completed each step in the activity, he was taught to move the picture from one column to another. In addition, visual representations of the step involved in requesting items from peers and getting his teacher’s attention were developed and reviewed before the math centers. Each day that Intervention 2 was implemented, Davis also practiced examples and non-examples of getting his teacher’s attention and requesting items from peers. Finally, Davis’ teachers were taught to use a hand signal to let him know how long he needed to wait before getting their attention. When he approached the teacher, she raised her hand and showed Davis five fingers. Then, she gradually lowered one finger at a time. When all five fingers were down, the teacher gave Davis her attention.

Gabe. In Intervention 1, Gabe was reminded of the behavioral expectations at the start of the activity. He was verbally praised when he was on-task. He also earned two class dollars for completing his work independently, or one class dollar for completing an assignment with assistance. When the target behavior occurred, Gabe was reminded he could raise his hand and to ask for help. If he did not complete the assignment during class, he was given the same assignment the next day. Inte-
vention 2 was identical but included teaching Gabe to use written instructions to complete his assignment independently. The teacher listed the steps Gabe needed to complete. He was then taught to cross off each step as he completed it.

Design

For each student, an A-B-A-C-A-[B-or-C] design was used to evaluate the effects of the different interventions. In Condition A (baseline), the teachers used the same procedures they had been using prior to the study. During Condition B, Intervention 1 was implemented. During Condition C, Intervention 2 was implemented. Sequentially, the initial baseline was followed by Intervention 1, a return to baseline, Intervention 2, another return to baseline, and then replication of the more effective intervention. Follow-up data were collected weekly for 3 weeks for Davis and Gabe, but could not be collected for Calli because the school year ended.

Data Collection and Analysis

The behavioral definitions used in Phase 1 were used again in Phase 2. Data on the replacement behavior (on-task) were collected daily. Treatment integrity (TI) data were collected during 33% of Calli’s sessions, 32% of Davis’ sessions, and 35% of Gabe’s sessions.

On-task behavior was measured using a 30-s whole-interval recording method. At the end of each interval, a plus was scored if the replacement behavior occurred throughout the entire interval. A minus was scored if the target behavior occurred at any time during the interval. Observations lasted 12 min for Calli and 15 min for Davis and Gabe.

Treatment integrity was assessed by using a checklist of the intervention procedures and recording whether each was implemented correctly. The level of treatment integrity was obtained by dividing the number of observed intervention procedures that were completed at the appropriate time by the total number of applicable intervention procedures, and multiplying the result by 100%.

Inter-Observer Agreement (IOA)

IOA data on student responding were collected by having a second observer independently record data on the replacement behavior. IOA data were collected in each condition and during 38% of Calli’s sessions, 32% of Davis’ sessions, and 30% of Gabe’s sessions. IOA was assessed using the exact interval-by-interval method (Kazdin, 1982). For on-task behavior, IOA averaged 93% (range = 88–100%) for Calli, 93% (range = 87–97%) for Davis, and 97% (range = 93–100%) for Gabe.

IOA for treatment integrity was established by having a second observer independently and simultaneously record whether each procedure was implemented correctly. IOA was calculated by dividing the number of agreements by the total number of applicable components and multiplying the result by 100%. IOA for treatment integrity averaged 100% for Calli and Davis, and 98% (range = 83–100%) for Gabe.

Social Validity

Social validity was assessed by using the Treatment Acceptability Rating Form-Revised TARF-R (Reimers & Wacker, 1988). This instrument includes 17 items, with multiple items addressing each of the following areas: reasonableness, effectiveness, side effects, disruptive-ness/time required, cost, and willingness. Each item is rated on a 7-point Likert-type scale. Scores can range from 17 to 119, with higher scores representing greater acceptability (Reimers & Wacker, 1988). The TARF-R was completed by each student’s teacher before implementing Intervention 1 and again before implementing Intervention 2.

Results

Figure 1 shows on-task levels in each condition for Calli (top panel), Davis (middle panel), and Gabe (bottom panel). Calli’s on-task behavior averaged 14% (range = 6–31%) during baseline, increased to 29% (range = 0–56%) with Intervention 1, and remained at an average of 29% (range = 20–44%) when Intervention 1 was withdrawn. Calli’s on-task levels increased to an average of 92% (range = 80–100%) when Intervention 2 was
implemented, decreased to an average of 51% (range = 0–92%) when Intervention 2 was withdrawn, and increased again to an average of 82% (range = 64–100%) when Intervention 2 was reinstated.

For Davis (middle panel), on-task levels averaged 47% (range = 43–50%) during baseline, increased to 62% when Intervention 1 was introduced, and decreased to an average of 31% (range = 20–50%) when the intervention was withdrawn. Davis’ on-task levels increased to an average of 89% (range = 83–93%) when Intervention 2 was introduced, decreased to a mean of 53% (range = 47–60%) when the intervention was withdrawn, and increased to an average of 89% (range = 77–100%) when Intervention 2 was reinstated. On-task levels remained high during follow-up.

Gabe’s on-task behavior averaged 11% (range = 0–20%) during baseline (bottom panel), increased to 18% (range = 0–37%)

Figure 1. Levels of on-task behavior (closed circles) and treatment integrity (open triangles) for Calli (top panel), Davis (middle panel), and Gabe (bottom panel).
with Intervention 1, and decreased to 3% (range = 0–10%) when Intervention 1 was withdrawn. Gabe’s on-task levels increased to an average of 83% (range = 73–93%) when Intervention 2 was introduced, decreased to an average of 10% when Intervention 2 was withdrawn, and increased again to 79% (range = 77–83%) when Intervention 2 was reinstated. During follow up, Gabe’s on-task behaviors maintained at an average of 94%.

Treatment integrity for Calli averaged 100% during intervention conditions and 0% (indicating no intervention implementation) during baseline conditions. For Davis, TI was 86% during intervention and 0% for baseline conditions. For Gabe, TI was 100% during intervention conditions, and 0% during baseline. Using the TARC-R, Calli’s teacher rated Intervention 1 with a score of 115 (out of 117) and Intervention 2 with a score of 117. Davis’ teacher rated Intervention 1 and Intervention 2 each with a score of 114. Gabe’s teacher rated Intervention 1 with a score of 111 and Intervention 2 with a score of 105.

Discussion

Some authors (e.g., Umbreit et al., 2007) have suggested that the design of function-based interventions must include consideration of whether an individual can independently perform the replacement behavior. Data from the present study support this contention. Individuals with established acquisition deficits were supported by two nearly identical interventions. When performance of the replacement behavior was taught directly, high levels of on-task behavior were produced. In contrast, when performance of the replacement behavior was prompted but not directly taught, on-task levels were barely higher than baseline.

Task analysis was used to determine whether each student could independently perform all components of the replacement behavior. The technique also made it possible to identify which particular steps in the task analysis of the replacement behavior required direct instruction. Although this procedure was used once before (Reeves et al., 2013), a reinforcement contingency was added in this study. The goal was to determine whether students could perform the replacement behavior, but would do so only under highly motivating conditions. Students who failed to perform the replacement behavior, even with contingent reinforcement, were deemed to have a clear acquisition deficit. A similar approach was used by Daly, Witt, Martens, and Dool (1997) to assess skill or performance deficits with academic tasks.

Each student’s level of on-task behavior dropped during the third reversal condition despite having received some instruction in how to perform the replacement behavior. This evidence suggests that, when initially establishing a new skill, it is necessary to maintain the environmental conditions that support that skill until the student performs it independently (i.e., fluently enough to be reinforced naturally). Depending on the complexity of the replacement behavior and the student’s ability to perform it, the amount of time required for instruction will vary and need to be monitored. Replacement behaviors will not maintain if intervention components are discontinued too quickly.

Social validity results indicated that the teachers did not differentiate between the two interventions. They found each to be helpful and acceptable, even though one was very effective and the other was not. Social validity data were collected before each intervention was implemented. As such, the rating of each intervention was not influenced by its effectiveness. In future work, staff should re-rate each intervention after using it.

Certain limitations should be noted. First, the function-based interventions were designed using the methods proposed by Umbreit et al. (2007). Different results may have obtained if a different intervention model or set of procedures had been used. Second, even though the task analyses were conducted by an experienced person following accepted procedures, the steps in each task analysis varied for each student. The data collection method used also varied. For Calli, a 30-s whole-interval method was used. For Davis and Gabe, data were collected on the number of steps each student performed independently. Task analysis and data collection procedures could vary from student to student, from practitioner to practitioner, and from replacement behavior to replacement behavior.

Third, because the study’s design required
frequently changing conditions, the interventions were primarily implemented by the first author. Although the teachers were informed about which condition was currently in effect for each student and the accompanying intervention components, they were not primarily responsible for implementing the intervention until the more effective intervention method was determined. This may have impacted the results. Finally, the instrument used to assess each student’s classroom environment (Ferro et al., 2008) has not been validated. The checklist did, however, provide data for systematically evaluating the classroom environment. Furthermore, when a second observer independently evaluated the same environment, high levels of IOA obtained.

The design of function-based interventions must include consideration of whether an individual can independently perform the replacement behavior. Failure to do so can result in interventions that are technically and procedurally accurate, yet quite ineffective. More attention to this important area is needed both in research and practice.

References


Received: 28 July 2016
Initial Acceptance: 27 September 2016
Final Acceptance: 17 November 2016
Concrete and App-Based Manipulatives to Support Students with Disabilities with Subtraction

Emily C. Bouck, Courtney Chamberlain, and Jiyoung Park
Michigan State University

Abstract: Manipulatives support students with and without disabilities in mathematics. However, as students age, concrete manipulatives can be limiting and potentially not age appropriate (Satsangi, 2015). An alternative is virtual manipulatives, including app-based manipulatives. This study compared the use of app-based manipulatives to concrete manipulatives in supporting students with disabilities in solving subtraction problems with regrouping. Using an adapted alternating treatment design with three middle school students with disabilities, the researcher found app base 10 blocks were more effective in terms of solving subtraction with regrouping for two of the students. They also found that all three students were more independent with the app-based manipulatives, although only two of the three students preferred the app-based manipulatives to the concrete manipulatives.

Students with disabilities struggle with mathematics, and generally more so than students without disabilities (The Nation’s Report Card, 2016). The average score for fourth-grade students with disabilities on the 2015 National Education Assessment Program (NAEP) was 218, as compared to the average score of 244 for fourth-grade students without disabilities. Similarly, for eighth-grade students with disabilities, the 2015 NAEP data suggested an average score of 247 for students with disabilities as compared to 287 for students without disabilities (The Nation’s Report Card, 2016).

Given the struggle students with disabilities face in mathematics, researchers and practitioners have sought effective practices. In mathematics, manipulatives are considered an effective strategy for teaching students with and without disabilities (Marley & Carbonneau, 2014). Mathematics manipulatives – generally assumed to be concrete manipulatives – are physical objects students can manipulate with their hands. Over the past few decades, an alternative form of mathematics manipulatives was developed and used to support students: virtual manipulatives. Virtual manipulatives are digital manipulatives that serve similar functions as concrete manipulatives – and are often similar to concrete manipulatives – but exist in a digital form (Bouck & Flanagan, 2010). Previously, virtual manipulatives were online (or Internet-based) manipulatives, such as ones available from the National Library of Virtual Manipulatives (NLVM). However, more recently, with the increase in attention to mobile devices, a newer form of virtual manipulatives exists: app-based manipulatives.

Although the use of concrete manipulatives is supported in research and practice for students with disabilities (Lai & Berkeley, 2012; Maccini & Gagnon, 2000), concrete manipulatives possess limitations. Hence, there is merit to considering virtual manipulatives – both online and app-based – for students with disabilities. For one, virtual manipulatives may be more age-appropriate for secondary students with disabilities (Satsangi, 2015). The use of concrete manipulatives typically designed for younger students (e.g., base 10 blocks) can be stigmatizing or embarrassing to use for secondary students. Virtual manipulatives may also reduce the cognitive load for students with disabilities (Suh & Moyer, 2008). Given the built-in supports or constraints within virtual manipulatives, students...
may experience a decreased cognitive load when using them (Moyer, Niezgoda, & Stanley, 2005; Suh & Moyer, 2008).

Previous Manipulatives Research

Prior research, albeit limited, exists on both the use of virtual manipulatives for students with disabilities and the comparison of concrete and virtual manipulatives for students with disabilities. In a multiple baseline single subject design study involving three high school students with learning disabilities, Satsangi and Bouck (2015) explored the use of online polynominoes (i.e., digital tiles) from the NLVM. The students successfully used the online manipulatives to solve area and perimeter problems as well as maintain and generalize these skills. Bouck, Flanagan, and Bouck (2015) also explored the use of online manipulatives to support middle school students with disabilities in solving area and perimeter problems. Also using the polynominoes from the NLVM, Bouck et al. (2015) conducted a pretest-posttest study and found that students answered more area and perimeter questions correctly – and attempted more problems – on the posttest following instruction with online manipulatives.

For comparison research, Bouck, Satsangi, Doughty, and Courtney (2014) compared concrete base 10 blocks to online base 10 blocks from the NLVM to support single-digit or double-digit subtraction with regrouping for three elementary students with autism. Bouck et al. (2014) found online and concrete manipulatives supported students correctly solving subtraction problems, although they were slightly more independent in completing the task analysis steps with the online manipulatives. The students also expressed a preference for virtual manipulatives. Finally, Satsangi, Bouck, Taber-Doughty, Bofferding, & Roberts (2016) compared the use of a online algebraic balance scale from the NLVM to a concrete algebraic balance scale in an alternating treatment design with three high school students with learning disabilities. Satsangi et al. (2016) found that the three students were successful with both types of manipulatives in solving linear algebra equations, although the students preferred the virtual manipulatives.

The existing research to date focuses on online manipulatives, such as from the NLVM. To date, very limited research exists on app-based manipulatives to support students with disabilities. The limited research is, of course, likely attributed to the relative recency of app-based manipulatives. Although not a manipulative app, recent research by Bryant et al. (2015) suggested no differences across three conditions were found in terms of students acquiring multiplication facts when comparing their learning via a math app, teacher-directed instruction, or a combination of app and teacher-directed instruction. Similarly, while not an app-based manipulative, Ok and Bryant (2015) found fifth-grade students with learning disabilities improved in their multiplication facts following an intervention of an app focused on multiplication fact practice.

While online manipulatives – such as the NLVM, provide educators with free Internet-based virtual manipulatives to support students across grades in multiple mathematical areas, online manipulatives require Internet access and do not work on tablets, such as iPads, given they often use JAVA. Given schools’ increasing use of mobile devices, such as iPads (Pilgrim, Bledsoe, & Riley, 2012), it is important to consider how app-based manipulatives compare to concrete manipulatives in supporting students in mathematics and if app-based manipulatives are a viable option for students with disabilities. The research questions for this study include: (a) what number of subtraction mathematics problems do students solve accurately when using app-based manipulatives and concrete manipulatives?; (b) what percentage of subtraction mathematics problems do students solve independently when using app-based manipulatives and concrete manipulatives?; and (c) what are student preferences when considering app-based or concrete manipulatives?

Method

Participants

Three middle students participated in the study. Each received special education services from their school district via a pullout pro-
gram (i.e., all core content courses taught by a special education teacher, including mathematics). Participants were chosen according to the following: (a) teacher recommendation for students struggling with double-digit or triple-digit addition or subtraction; (b) confirmation of struggles with double- or triple-digit addition or subtraction through independent KeyMath™-3 Addition and Subtraction subtest with ceilings in double- or triple-digit addition and subtraction, and (c) fine motor ability to move concrete manipulative blocks and navigate a touch-based iPad application.

José. José was a 14-year-old Hispanic student in eighth grade. José was a pleasant young man who played on his school’s basketball team and strongly disliked math. His special education eligibility was in the area of mild intellectual disability. According to his performance on the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV), José’s full-scale IQ was 62. His performance on the mathematics subtests of the Wechsler Individual Achievement Test-Third Edition (WIAT-III) indicated his Math Problem Solving standard score of 71 and Numerical Operations standard score of 66 were in the below average and low ranges of performance, respectively. Additional mathematics scores from the KeyMath™-3 Diagnostic Assessment and STAR Math Assessment suggested grade equivalent mathematics performance at the second grade level (2.9 and 2.1, respectively). José’s teacher indicated his mathematics performance was inconsistent, especially with regards to triple-digit subtraction.

Ellen. Ellen was a 13-year-old Caucasian student in seventh grade. Ellen was a friendly girl who enjoyed drawing and coloring. Her special education eligibility was in the area of mild intellectual disability. According to her performance on the WISC-IV, Ellen’s full-scale IQ was 68. Her standard score of 62 on the Woodcock Johnson Tests of Achievement-Third Edition (WJ-III-Ach) Broad Math composite was in the very low range of performance. Her score on the STAR Math Assessment suggested grade equivalent mathematics performance at the third grade level (3.2). Ellen also had mental health issues and took medication for bipolar disorder. She had inconsistent attendance and often complained of not feeling well.

Vince. Vince was an 11-year-old Hispanic male in sixth grade. Vince was very quiet and reserved but always prepared for class and for sessions. His special education eligibility was in the area of specific learning disability for reading comprehension and mathematics calculation. According to his performance on the Woodcock-Johnson Tests of Cognitive Abilities-Third Edition (WJ-III-Cog), Vince’s Global Intellectual Ability score was 84. His standard score of 79 on the Applied Problems subtest and 73 on the Calculation subtest of the WJ-III-Ach was in the low range of performance. Additionally, his score on the KeyMath™-3 Diagnostic Assessment suggested grade equivalent mathematics performance at the second grade level (2.2), while his score on the STAR Math Assessment suggested grade equivalent performance at the third grade level (3.3). Vince struggled with processing, as evident by both researcher observation as well as stated by his special education teacher. When asked a question or tasked with solving a problem with the manipulatives, he was very slow to respond. His teacher also discussed his processing struggles and his struggle to apply concepts repeatedly presented.

Setting

Study sessions occurred at a public middle school in a rural Midwest town. At the time of data collection, the school enrolled approximately 712 students in sixth through eighth grade. Approximately 86% of the student population were Caucasian, 12% Hispanic, 1% Multiracial, and less than 1% were African American and Alaska Native/American Indian. Approximately 16% of the student body were identified as students with disabilities. Data collection occurred in a small room in the building’s administrative area as well as occasionally unoccupied areas of the school library. Each space had at least three chairs and tables large enough for the students to use the concrete manipulative blocks or iPad as well as write their responses down on probes. Sessions were conducted in an one-on-one environment between a researcher (one of the authors) who implemented the study and the student; sessions in which inter-
observer agreement data were collected had a second researcher present.

Materials

Materials included pencils, researcher-constructed probe sheets, concrete manipulative blocks and place value sheets, and an iPad with the base 10 blocks app. The researchers gave students a pencil in each session and a probe sheet containing five subtraction problems with regrouping, with each problem presented vertically. José and Vince completed assessments with triple-digit subtraction problems while Ellen double-digit subtraction problems, selected based on students’ instructional mathematics level and performance on the KeyMathTM-3. Note, each five-question problem set was unique. The researchers developed all double-digit and triple-digit subtraction problems needed across the study and then randomly selected problems for each problem set; each problem set was randomly selected for its use (i.e., phase of the study).

During the concrete manipulative blocks phase, students were provided with a place value sheet and the concrete base 10 blocks. The place value sheet contained three columns with pictorial representations of manipulative blocks at the top of each column. The right column was labeled “Ones” with a ones block at top, the middle column was labeled “Tens” with a tens block at top, and the left column was labeled “Hundreds” with a hundreds block at top. Two sheets were provided, one for the minuend and one for the subtrahend; the place value sheets were also consistent with the virtual manipulatives app. Researchers provided two different colored base 10 blocks; students used one color for the minuend and the other for the subtrahend. Researchers provided more than enough ones, tens, and hundreds blocks – as applicable – at the start of each session to the left side of the place value sheets. Students set up blocks on the place value sheets to represent each number of the problem, regrouping tens or hundreds blocks for ones or tens when regrouping was required to solve the problem, and removing blocks when subtracting.

During the app-based manipulative blocks phase, students were provided with an iPad on which a manipulative app was downloaded. The app, Base 10 Blocks Manipulative Version 1.1.0, was developed by Brainingcamp, LLC and available for a $0.99 purchase on iTunes. The researchers selected this app after evaluating multiple base 10 block app for iOS devices. Although the app has limitations, this app possesses many positives to make it an appropriate choice. For one, the blocks on the app – and how they operate – are very similar to concrete base 10 blocks. The app was also flexible to support students in place value, addition, and subtraction, and worked with multiple place values, including decimals. Finally, the app from Brainingcamp was intuitive to use but also had built-in constraints beyond what concrete manipulatives could do (i.e., it would not let students regroup from the subtrahend rather than the minuend).

Depending on the type of subtraction problem, the app presented two to three columns and a separate row for both the minuend and the subtrahend. As with the place value sheets in the concrete manipulative blocks phase, the top of the columns contained pictorial representations of ones, tens, and hundreds, if applicable, blocks. Students set up blocks by touching the pictorial block representations and dragging them from the tops of the columns to the respective rows for the minuend and the subtrahend. To ungroup a hundreds or tens blocks, students moved the hundreds blocks to the tens place and/or tens blocks to the ones place. When this was done, the block being moved ungrouped itself into smaller units (i.e., ten tens blocks or ten ones blocks). To subtract, students dragged blocks from the subtrahend row to same-unit blocks in the minuend row. When two same-unit blocks from each number came into contact, they became semi-transparent to signify subtraction (i.e., as if the student had removed the blocks from the place value chart). Of note, the iPad app displayed the problem – and the answer – on the app at the bottom of the screen. Researchers covered the problem and answer with small sticky notes; no student ever tried to remove the sticky notes from the iPad or questioned researchers about it.
Independent and Dependent Variables

The independent variable for the study was use of manipulatives; students used concrete manipulative blocks or app-based manipulative blocks to complete subtraction problems. Using concrete manipulative was defined as setting up and moving the base 10 blocks to solve each subtraction problem. Using an app-based manipulative was defined as setting up and dragging the digital base 10 blocks in the app to solve each subtraction problem.

The dependent variables for the study included (a) the number of subtraction problems the student answered correctly out of five (i.e., accuracy), (b) the amount of time it took students to complete each assessment (i.e., task completion time), and (c) the percentage of subtraction task analysis steps they completed independently without prompting (i.e., independence). The researchers used event recording for accuracy and independence and duration for task completion time. Accuracy was calculated by summing the number of problems a student answered correctly on a probe out of five. Task completion time was measured by the amount of time it took a student to solve the five problems, with or without manipulatives; the timer was started when the student received the subtraction probe sheet and ended when the student wrote the last answer on the probe. Independence was calculated by determining the number of task analysis steps the student completed independently in each session and dividing that by the total number of steps (concrete, app-based, and no manipulative task analysis recording sheets available upon request for double-digit and triple-digit subtraction). There were 14 steps for the triple-digit subtraction problems with both manipulative (total of 70 across each probe), and 9 steps for double-digit subtraction problems with manipulative, for a total of 45.

Experimental Design

An adapted alternating treatment design was employed, with four phases: baseline, intervention, best treatment, and generalization (Sindelar, Rosenberg, & Wilson, 1985; Wolery, Gast, & Hammond, 2010). In the present study, students solved subtraction problems across three alternating conditions during the intervention phase: concrete manipulatives, app-based manipulatives, and extended baseline in which no manipulatives were used. Five sessions of each condition were alternated at random, with no more than two sessions of the same condition (concrete, app, or no manipulative) in a row. Using this design, the authors were able to determine the effectiveness of each type of manipulative on the students’ subtraction problem solving and create an experimental control for each student (Sindelar et al., 1985). Across all sessions, the researchers served as the implementers; the teacher was not involved with the implementation of the study procedures and only participated in social validity.

Procedures

Baseline. For baseline, students were required to complete subtraction assessments at their mathematics level (i.e., triple-digit or double-digit subtraction with regrouping) with no manipulatives. Each subtraction probe was a sheet of paper with five subtraction problems presented in two columns on one side of the sheet, and each problem was presented vertically. Students were provided with a pencil and asked to solve the five problems. Researchers were prepared to offer prompting if students failed to initiate solving within 10 seconds.

Pre-training. Before intervention, students were provided training on both the concrete manipulatives and app-based manipulatives. To train students on each type of manipulative, the researchers employed explicit instruction (Doabler & Fien, 2013). With the explicit instruction, the researchers first demonstrated and used think-alouds for how to use the manipulatives to solve practice subtraction problems at their mathematics level (i.e., double- or triple-digit subtraction problems with regrouping); the modeling portion for the explicit instruction was done for two problems each session. Consistent with explicit instruction, after two sessions of modeling (i.e., demonstrating and use think-alouds), the students worked to solve two problems, and researchers provided prompts and cues as needed. Finally, after two sessions with guiding, the students completed five
problems independently. If students correctly solved 80% of the problems in the independent phase with each manipulative, they were considered successfully trained with that manipulative. Each student had to score 80% during one training probe for each manipulative type to move into the intervention phase. Ellen was trained in one session for each manipulative condition. José required two sessions for each manipulative condition. Vince was trained after three separate sessions for both the concrete and app-based manipulatives.

During the concrete manipulatives explicit instruction, students were trained to read the subtraction problem on the probe sheet and set up the hundreds blocks (triple-digit subtraction only), tens blocks, and ones blocks for the minuend in the first row of the place value sheet using one of the two sets of different color concrete manipulative blocks. These steps were repeated for subtrahend using the remaining set of concrete manipulative blocks, and students were trained to place blocks for the subtrahend in the second row of the place value sheet. If the minuend ones were smaller than the subtrahend ones, students were trained to regroup a tens block from the minuend for 10 ones blocks. The ten ones blocks were then added to the ones of the minuend. Students solving triple-digit problems were also trained to regroup a hundreds block from the minuend into 10 tens blocks should the minuend tens be smaller than those in the subtrahend. The 10 tens blocks were then added to the tens of the minuend. Finally, students were trained to subtract subtrahend blocks from same-unit minuend blocks and write the answer on the probe.

Similarly, students were trained to solve subtraction problems using the app-based manipulatives. As with the place value sheets for the concrete manipulatives, the app displayed two rows in which students were trained to drag ones blocks, tens blocks, and hundreds blocks (for triple-digit subtraction only) to set up the minuend and subtrahend subtraction problem numbers. Minuend blocks were green, while subtrahend ones were red. If students accidentally selected too many blocks, they dragged the extra blocks to a small trash bin icon in the app to delete. Next, students were trained to regroup tens and hundreds blocks in the minuend by dragging them to the column to the right (i.e., the ones and tens column, respectively). Upon doing so, the blocks ungrouped into smaller units (e.g., 10 ones blocks from the ten’s block and 10 tens blocks from hundred’s block). To subtract, students were trained to drag blocks from the subtrahend to same-unit blocks in the minuend. When paired with a same-unit block, the paired blocks became semi-transparent, indicating they were subtracted. Once all the blocks from the subtrahend were paired with blocks from the minuend, or subtracted, they were trained to count the remaining opaque blocks and write the answer on the probe.

Intervention. Following trainings, students completed the intervention phase, alternating between concrete manipulative blocks, app-based manipulative blocks, and no manipulatives (i.e., extended baseline) to solve subtraction problems with regrouping. Each condition included five sessions, with five subtraction problems per session. Session order was randomly determined with no more than two of the same conditions occurring consecutively. The system of least prompts was used during the intervention phase to assist students if they did not engage in the correct task analysis step within 10 seconds of completing the previous step (i.e., system of least prompts delivered after a 10-second time delay). The researchers used prompts that ranged from less intense visual (i.e., gesturing) and verbal prompts (i.e., indirect verbal prompts such as, “What comes next?”) to more intense verbal prompts (i.e., direct verbal prompts indicating the next step), modeling, and partial physical assistance (Doyle, Wolery, Ault, & Gast, 1988). More intensive prompts were given when students did not respond to previous prompts.

Concrete manipulative blocks. During the concrete manipulative intervention condition, students were provided with two place value sheets, two different colored sets of concrete base 10 blocks, a probe sheet with five subtraction questions, and a pencil. The two sets of blocks were placed on the table for students to select from for the subtraction problem minuend and subtrahend. Upon reading the subtraction problem on the probe sheet, students were expected to set up the correct number of
blocks in each cell of the place value sheets. They also regrouped larger blocks for smaller ones in the minuend in order to subtract the subtrahend from the minuend. Next, they counted the remaining blocks and wrote the answer on the probe sheet.

**App-based manipulative blocks.** In this condition, students used the Base 10 Blocks Manipulative iPad app (Brainingcamp, LLC, 2015) to solve subtraction problems. In addition to the iPad, students were provided with a probe sheet with five subtraction problems and a pencil. The iPad was opened to the app, and researchers adjusted the app settings prior to the session to fit students' mathematics levels. The app for Ellen had two rows and columns to represent double-digit subtraction, while the app for José and Vince had two rows with three columns to represent triple-digit subtraction. Students touched and dragged blocks down from representative pictures at the top of each column to represent the problem. After the blocks were set up, students regrouped and subtracted. Next, students counted the remaining blocks and wrote the corresponding answer on the probe sheet. As noted previously, sticky notes were placed over the problem and answer at the bottom of the iPad display so students could not see them.

**No manipulative.** The baseline condition of no manipulatives was extended into the intervention phase to serve as a control for student performance to determine a functional relation. Researchers provided students with a five problem probe sheet and a pencil. Students solved the problems without any aids (i.e., concrete or app-based manipulatives).

**Best treatment.** Following the intervention phase, students completed three sessions using the intervention condition – concrete or app-based manipulatives – most effective for them during intervention. To determine best treatment, researchers used percentage of non-overlapping data (PND; Gast & Spriggs, 2010); PND is a commonly used method within alternating treatment designs to compare the data from one condition to another (Gast & Spriggs, 2010). Note, PND was not used an effect size for accuracy; the researchers used Tau-U for calculate accuracy effect size, as will be explained later. Specifically, the researchers determined PND first for the accuracy, and then, if there was no difference, for the independence. To calculate PND, the researchers found the number of sessions that one condition (e.g., app-based manipulatives) was more effective than the other (e.g., concrete manipulatives) in the target dependent variable – first by accuracy and then independence, if needed. They then divided the summed number by five and multiplied by 100 (Wolery, Gast, & Hammond, 2010). For Ellen, best treatment was calculated based on independence data as both manipulative conditions were equally effective for her, while for José and Vince, best treatment was calculated based on accuracy data. For all three students, the best treatment condition corresponded to their preferred condition.

Similar to the intervention phase, students were provided with a probe sheet with five problems and a pencil in addition to the concrete manipulative blocks or iPad with the app, depending on the condition most effective for him or her. Best treatment followed the same procedures as intervention and included the same data collection (i.e., accuracy, task completion time, and independence). The system of least prompts was also used when students did not initiate a step or the correct step within 10 seconds of completing the previous one.

**Generalization.** Two generalization sessions were implemented following best treatment to evaluate for any generalization of the subtraction skills. Consistent with baseline and extended baseline, students were given a five problem probe sheet and a pencil. No aids in the form of concrete or app-based manipulative blocks were provided to assist students with problem solving.

**Inter-Observer Agreement and Treatment Integrity**

Inter-observer agreement (IOA) data and treatment integrity data were recorded for both the independent variable (manipulative condition) and prompts (i.e., independence data). IOA data were recorded for each student for (a) two sessions (40%) for each intervention condition (concrete manipulative, app, or no manipulatives), (b) at least 40% of baseline sessions (two for Ellen and Vince and three for José), (c) 33% for the best treatment sessions, and (d) 50% of the generalization.
sessions. During sessions in which IOA data were collected, two researchers were present and collected data. IOA was calculated by summing the number of agreements for both dependent variables examined – accuracy and independence – and dividing it by the number of agreements plus disagreements. IOA was 100% for accuracy for each student for each phase and condition. IOA for independence was 100% for all three students during baseline, extended baseline during intervention, and generalization. It was also 100% for José and Ellen for both intervention conditions and best treatment; IOA for Vince was 97.2% for the app intervention condition and best treatment and 92.9% for the concrete manipulatives condition.

Treatment integrity data were recorded across 40% of intervention and 33% of generalization sessions. The integrity data monitored included whether the students were provided a probe sheet, the appropriate type of manipulatives, and researchers implemented the system of least prompts. Treatment fidelity for each student for each condition and phase was 100%.

Social Validity

Social validity interviews were conducted after intervention with the teacher and the students. Students were asked questions regarding their perceptions about each type of manipulatives, including which type they preferred. The students’ special education teacher was interviewed about her mathematics instructional practices, her students’ mathematics abilities, and her opinion of using each type of manipulative during instruction.

Data Analysis

To analyze the data, researchers conducted visual analysis, including calculating level, trend, and effect sizes (Gast & Spriggs, 2010). Researchers calculated level by finding the stability of the data in each phase or condition. To do this, researchers first calculated the median for each dependent variable for each student, and then a 20% interval around each median, referred to as stability envelope (e.g., if the median was 4, the stability envelope range would be 3.2–4.8). Data were stable if 80% of a student’s data fell within 20% of the median for each dependent variable analyzed (Gast & Spriggs, 2010). Researchers used the split-middle method to determine trend (White & Haring, 1980). Researchers split the data for each phase and condition in half and calculated the mid-rate and mid-date. Next, they drew a line between the mid-rate and mid-date and determined if the line was accelerating, decelerating, or zero-celerating (Gast & Spriggs, 2010). The researchers used Tau-U to calculate effect size for the accuracy data. Tau-U contrasts each intervention condition with baseline (Parker, Vannest, Davis, & Sauber, 2011). To calculate the Tau-U, researchers used the web-based calculator (see http://www.singlecaseresearch.org/calculators/tau-u; Vannest, Parker, & Gonen, 2011). Tau-U scores less than or equal to 65% suggest a small effect, 66–92% a medium effect, and 93% and above a large effect (Parker, Vannest, & Brown, 2009).

Results

José

José struggled with triple digit subtraction with regrouping. José’s accuracy with the app base 10 blocks was superior to that with the concrete base 10 blocks with a PND of 40% (concrete to app PND = 20%). José also indicated his preferred treatment condition was the app.

José’s accuracy data. During baseline, José’s average number of problems correct was 2, with a range from 0–4 (see Table 1 & Figure 1). José’s inconsistent performance during baseline was consistent with teacher reports. During intervention, he averaged 3.8 and 4.0 problems correct in the concrete and app conditions, respectively. Compared to baseline data, José’s Tau-U for accuracy was 69% for concrete and 83% for app-based manipulatives – both a medium to high effect. José’s extended baseline showed slight improvement from baseline ($\mu = 2.4$ vs. $\mu = 2$), although he answered more correctly, on average, during generalization (3.5). José’s average score during best treatment was higher than during his intervention condition ($\mu = 4.3$ vs. $\mu = 4$).
<table>
<thead>
<tr>
<th>Participant</th>
<th>DV</th>
<th>Data</th>
<th>Baseline</th>
<th>Concrete Manipulative</th>
<th>App-based Manipulative</th>
<th>No Manipulative</th>
<th>Best Treatment</th>
<th>Generalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>José</td>
<td>Accuracy</td>
<td>μ</td>
<td>2</td>
<td>3.8</td>
<td>4</td>
<td>2.4</td>
<td>4.3</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>variable</td>
<td>variable</td>
<td>variable</td>
<td>variable</td>
<td>variable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>decelerating</td>
<td>accelerating</td>
<td>accelerating</td>
<td>accelerating</td>
<td>accelerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Independence</td>
<td>μ</td>
<td>100%</td>
<td>97.5%</td>
<td>98.2%</td>
<td>100%</td>
<td>100%</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>stable</td>
<td>decelerating</td>
<td>decelerating</td>
<td>decelerating</td>
<td>stable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>zero-celerating</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>μ</td>
<td>3.4</td>
<td>13.5</td>
<td>8.7</td>
<td>2.9</td>
<td>7.8</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>stable</td>
<td>decelerating</td>
<td>stable</td>
<td>decelerating</td>
<td>stable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>decelerating</td>
<td>accelerating</td>
<td>accelerating</td>
<td>decelerating</td>
<td>decelerating</td>
<td></td>
</tr>
<tr>
<td>Ellen</td>
<td>Accuracy</td>
<td>μ</td>
<td>0.4</td>
<td>5</td>
<td>5</td>
<td>1.8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>stable</td>
<td>stable</td>
<td>stable</td>
<td>variable</td>
<td>stable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Independence</td>
<td>μ</td>
<td>100%</td>
<td>98.2%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>stable</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>zero-celerating</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>μ</td>
<td>2.7</td>
<td>11.1</td>
<td>8.2</td>
<td>2.4</td>
<td>7.2</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>variable</td>
<td>variable</td>
<td>stable</td>
<td>variable</td>
<td>variable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>decelerating</td>
<td>decelerating</td>
<td>accelerating</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td></td>
</tr>
<tr>
<td>Vince</td>
<td>Accuracy</td>
<td>μ</td>
<td>0.4</td>
<td>4.2</td>
<td>4</td>
<td>0</td>
<td>4.7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>variable</td>
<td>stable</td>
<td>variable</td>
<td>variable</td>
<td>stable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>accelerating</td>
<td>decelerating</td>
<td>zero-celerating</td>
<td>zero-celerating</td>
<td>accelerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Independence</td>
<td>μ</td>
<td>100%</td>
<td>88.0%</td>
<td>91.5%</td>
<td>100%</td>
<td>94.8%</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>stable</td>
<td>deccelerating</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td>accelerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>zero-celerating</td>
<td>deccelerating</td>
<td>accelerating</td>
<td>zero-celerating</td>
<td>accelerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>μ</td>
<td>2.8</td>
<td>23.4</td>
<td>19.6</td>
<td>3.2</td>
<td>21.7</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>level</td>
<td>variable</td>
<td>stable</td>
<td>variable</td>
<td>variable</td>
<td>accelerating</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trend</td>
<td>decelerating</td>
<td>decelerating</td>
<td>decelerating</td>
<td>accelerating</td>
<td>decelerating</td>
<td></td>
</tr>
</tbody>
</table>
Jose’s independence data. Jose was very independent during all phases of the study, but was 100% independent during baseline, extended baseline (no manipulative condition during the intervention phase), best treatment, and generalization. Across the intervention conditions, Jose needed few prompts, with an average of 97.5% independence across the task analysis steps in the concrete manipulative condition and 98.2% independence across the task analysis steps in the app-based manipulative condition (see Figure 2). When Jose needed prompts, as few as they were, they were gestures and usually involving steps related to regrouping. Jose proceeded through steps so quickly at times that he would just skip over a step in an effort to finish quicker.

Jose’s task completion time data. Jose completed the five problems during baseline, extended baseline, and generalization phases in a relatively short amount of time – an average of 3.4 minutes, 2.9 minutes, and 3.3 minutes, respectively (see Figure 3). In terms of efficiency, Jose spent less than one minute per problem during baseline, extended baseline, and generalization. Jose completed the five problems in a shorter amount of time in the app-based manipulative condition (µ = 13.5 minutes) as compared to the concrete manipulative condition (µ = 8.7 minutes). The difference in efficiency (minutes spent per problem) when comparing the concrete and app-based averages was about one minute (2.7 minutes per problem to 1.7 minutes per prob-
Jose decreased his task completion rate during best treatment ($u = 7.8$ minutes).

Ellen's data indicated difficulty with double-digit subtraction with regrouping. Ellen's best treatment was selected as app-based manipulatives. Although she was equally effective with both app and concrete manipulatives with correct digits (100%), she was slightly more independent with the app manipulatives (PND 60%). Ellen also preferred the app manipulatives.

Ellen's accuracy data. During baseline, Ellen's average number of problems correct was 0.4, with a range from 0–2 (refer to Table 1). She answered 100% of the double-digit problems correct in both intervention conditions as well as during best treatment. Compared to baseline data, Ellen’s Tau-U for accuracy was 100% for both concrete and app-based manipulatives – a large effect. While Ellen's extended baseline showed slight improvement from baseline ($u = 1.8$ vs. $u = 0.4$), she answered all problems correct during both generalization probes. Ellen's accuracy data for both intervention conditions were stable and zero-celerating.

Ellen's independence data. Like José, Ellen was very independent during all phases of the study, including 100% during baseline, extended baseline (no manipulative condition during the intervention phase), best treatment, and generalization. Across the intervention conditions, Ellen needed few prompts, and never more than a gesture. Specifically, she averaged 98.2% independence across the task analysis steps in the concrete manipulatives condition and 99.4% independence across the task analysis steps in the app-based manipulative condition. Ellen’s independence data for both intervention conditions were stable; her data had a zero-celeration trend for the app-based manipulatives and an accelerating trend for the concrete manipulatives.

Ellen's task completion time data. Ellen completed the five subtraction problems during baseline, extended baseline (no manipulative), and generalization phases in less than three minutes. Ellen was faster in solving the five problems with the app-based manipulatives ($u = 8.2$ minutes) as compared to the concrete manipulatives ($u = 11.1$ minutes). In terms of efficiency, Ellen took, on average, 1.6 minutes per problem with the app-based manipulatives and 2.2 with the concrete manipulatives. Her task completion time during best treatment with the app-based manipulatives was less than during intervention ($u = 7.2$ minutes), indicating greater efficiency.

Vince

Like José, Vince struggled with triple-digit subtraction with regrouping. Unlike Ellen and José, though, Vince was more accurate with the concrete manipulatives (PND 40% as compared to a PND of app to concrete of 20%).
Vince indicated he preferred concrete manipulatives.

**Vince’s accuracy data.** Vince averaged 0.4 problems correct during baseline, with a range of 0–1 (refer to Table 1). His average number of problems correct was 4.2 with concrete manipulatives and 4 with app-based manipulatives. Compared to baseline data, Vince’s Tau-U score was 100% for both concrete and app-based manipulatives—a large effect. Vince answered zero problems correctly during either extended baseline or generalization. He was most accurate during best treatment ($\mu = 4.7$, range 4–5). Vince’s data for concrete manipulatives was stable with an accelerating trend; they were decelerating and unstable for the app-based manipulatives.

**Vince’s independence data.** Vince struggled to engage independently in the task analysis steps with both the concrete and app-based manipulatives, although he required fewer steps prompted with the app-based manipulatives ($y = 91.5\%$ vs. $88.0\%$). He needed multiple levels of prompting, including gesture, indirect verbal, direct verbal, and modeling. Vince’s most prompted task analysis steps involved regrouping hundreds or tens blocks. His independence data for both conditions were stable; his app-based manipulative independence data had an accelerating trend while he had a decelerating trend for the concrete manipulatives.

**Vince’s task completion time data.** Vince completed the five subtraction problems during baseline, extended baseline (no manipulative), and generalization phases in less than four minutes. However, he averaged around or over 20 minutes for the two manipulative conditions ($\mu = 23.4$ minutes for concrete and $\mu = 19.6$ minutes for app). In terms of efficiency, Vince took, on average, 4.7 minutes per problem with the concrete manipulatives and 3.9 with the app-based manipulatives. His average task completion time during best treatment with the concrete manipulatives was less than during intervention ($\mu = 21.7$ minutes).

**Social Validity**

Ellen and José both indicated they preferred the app-based manipulatives, stating they felt they were easier to use despite that they had only used concrete manipulatives previously. Vince, however, preferred the concrete manipulative and indicated he was more comfortable with them. Vince and José expressed that they were both excited for the study to be over because neither enjoyed doing math, while Ellen liked math and liked to work one-on-one with the researchers. The teacher was excited about the potential of the app-based base 10 blocks and was going to look into seeing if it might be able to be purchased and additional iPads secured for her students. She also stated that she wished at least one other student returned a consent form as she felt he too could have benefited from additional one-on-one work with subtraction.

**Discussion**

This study explored how app-based manipulatives compared to concrete manipulatives in supporting students with disabilities in solving subtraction problems with regrouping. To compare both manipulatives in terms of accuracy, independence, and task completion time, researchers used a single subject adapted alternating treatment design with three middle school students—two with mild intellectual disability and one with a learning disability. The main result is that app-based base 10 blocks can be just as effective and efficient as concrete base 10 blocks.

Although the results were idiosyncratic in terms of what type of manipulative was more effective, all three students improved in terms of correctly solving the problems during the intervention conditions as compared to baseline. These results are consistent with previous research that suggests manipulatives—both concrete and app-based—are effective for students with disabilities (c.f., Bouck et al., 2014; Satsangi & Bouck, 2015). In addition, both types of manipulatives were able to support students; there was not a clear, consistent difference in terms of accuracy between the two types of manipulatives. In other words, it could be suggested that app-based manipulatives could equally support students with disabilities as compared to concrete manipulatives, also consistent with previous research (Bouck et al., 2014; Satsangi et al., 2016).

Despite prompting and use of manipulatives, Vince and José did not always achieve
100% accuracy during the intervention conditions. While the researchers prompted for students to set up the problems correctly in terms of following the task analysis steps, if students made a subtraction error, they were not prompted to correct; Vince and José made subtraction errors. With the concrete manipulatives, errors involved removing too many ones blocks during the subtraction, and with the app-based manipulatives, particularly with Vince, the errors involved inherent elements within the app. Two issues were at play. One was when subtracting with the app-based manipulatives, the blocks would become semi-transparent (or ghost out), and Vince would struggle to see all the opaque blocks. The other issue involved the small screen size; when there were 15 tens blocks, for example, they could be difficult to see or even hidden and sometimes they would be missed in the counting after the subtraction occurred.

In terms of independence, students needed more prompts during intervention as compared to baseline and extended baseline (i.e., no manipulatives). However, even though students answered some problems correctly during baseline and extended baseline, they were not consistently accurate. Overall, Ellen and José needed few prompts during the intervention conditions and zero during their best treatment of app-based manipulatives. When prompts were given, it was almost always a gesture and related to regrouping. Regrouping during subtraction was a challenge for all three students, and, when analyzing their digits correct across all phases, the reason they were largely incorrect without the manipulatives. Regrouping is a challenging concept for students, including many students with disabilities (Witzel, Ferguson, & Mink, 2012). There is evidence from the accuracy scores during generalization and, to some extent, extended baseline that suggest Ellen and José were beginning to internalize the regrouping step in their subtraction problems even when manipulatives were not present. Vince, on the other hand, needed more prompting. His performance was consistent with his processing challenges. The researchers also found that the 10-second delay reduced the numbers of prompts that might otherwise be given, as he often performed a step at 8 or 9 seconds within the time delay.

Students were more independent in solving the problems during baseline; they all experienced an increased amount of prompting – or conversely a decrease in independence – during intervention as compared to baseline. Although minimal for Ellen and José, one hypothesis is that students may have needed more training with both types of manipulatives prior to entering the intervention phase. However, Ellen and José both experienced the slight increase in prompting during the middle of the intervention period for both manipulatives, rather than the beginning. These results suggest that perhaps if the intervention sessions occurred more frequently (i.e., more than twice a week or without breaks due to absences), Ellen and José would have been more independent. The more training hypothesis could be appropriate for Vince, given that he needs more prompting for the app-based manipulative during the first session.

All three students were more independent with the app-based manipulatives. One hypothesis for these data is that the virtual – or app-based – manipulatives have built in constraints or structure to further support the student. For example, with the app-based manipulatives, one cannot regroup a tens block from the subtrahend as it is prohibited within the app. With the concrete manipulatives, Vince more than once attempted to regroup from the subtrahend rather than the minuend and the researchers would need to prompt. Hence, any form of virtual manipulatives may work to reduce students’ cognitive load (Suh & Moyer, 2008).

Based on the accuracy, independence, and task completion time data, the researchers hypothesize that Ellen and José could be supported in learning subtraction with regrouping through the Concrete-Representational-Abstract (CRA) approach; the CRA approach is considered an evidence-based or effective approach for teaching mathematical concepts to students with disabilities (Agrawal & Morin, 2016). The researchers further hypothesize that both Ellen and José could be taught the CRA approach with app-based manipulatives substituted in place of concrete. Not only was this deemed to be the best treatment condition for
both students, but it was also their preferred condition. Ellen and José both indicated a preference for app-based manipulatives, given their perception that they were easier and, as supported in other studies, more age appropriate and less socially stigmatizing (Satsangi & Bouck, 2015).

**Implications for Practice**

This study suggests implications for practice. For one, it suggests students with disabilities – such as mild intellectual disability and learning disability – can be supported in solving mathematical problems with app-based manipulatives. The use of an alternative to concrete manipulatives – while beneficial – is particularly important for older children, given that many manipulatives may not be age appropriate for secondary students and undesirable socially (e.g., concrete base 10 blocks). App-based manipulatives allow older students to be supported in a less conspicuous manner, and hence, students may be more likely to use the tool. Second, it does suggest that app-based and concrete manipulatives may be interchangeable and thus teachers implementing the CRA approach for students with disabilities or struggling students may be able to use app-based manipulatives in place of concrete within the sequence.

**Limitations and Future Directions**

This study has multiple limitations. With the adapted alternating treatment design was used, solving the math problems with concrete and app-based manipulatives were not probed during baseline. The researchers only probed for no manipulatives during baseline and extended it throughout the intervention as the control set (Wolery et al., 2010). One reason for not probing the manipulative conditions during baseline was the repeated exposure students had previously to the concrete manipulatives. However, the differential amount of exposure (i.e., they had never used app-based manipulatives prior to the study while the teacher routinely used concrete manipulatives) could be considered another limitation. In addition, a pre-determined set of criteria was not used to determine the intervention phase (Wolery et al., 2010). Another limitation involves that the researchers implemented the system of least prompts along with the intervention conditions; the impact of the system of least prompts was not separated. A final limitation involves that Vince, in particular, may have benefited from additional training prior to entering the intervention, especially with the app-based manipulatives.

In terms of future research, researchers should continue to explore the use of app-based manipulatives to support students with disabilities in mathematics, including conducting group design research. Researchers should also extend the exploration of app-based manipulatives to other mathematical areas, such as area and perimeter and algebra. In addition, given the few minor concerns regarding the actual app selected, researchers should explore other app-based base 10 manipulatives as their effectiveness and efficiency in supporting students with disabilities in subtraction, or even addition or multiplication. Finally, researchers should examine the use of virtual – including app-based and online – manipulatives within the approach, given the effectiveness of CRA in supporting students with disabilities.

**References**


Brainingcamp, LLC. (2016). Base Ten Blocks Manipulative (Version 1.1.0) [Mobile application software].

Bryant, B. R., Ok, M., Kang, E. Y., Kim, M. K., Lang,


Concrete vs. App Manipulatives / 331
An Analysis of State Guidelines for Intellectual Disability

Edward A. Polloway and Maryleen Auguste
University of Lynchburg

J. David Smith
University of North Carolina-Greensboro

Delia Peters
University of Lynchburg

Abstract: This study provides an updated analysis of state guidelines with respect to terminology and definitions in the field of intellectual disability (ID). The study serves as a methodological replication of prior work that has been reported in the literature in earlier decades. The data were acquired by reviewing web-based state guidelines for ID terminology and definition. These data are reported for all 50 states plus the District of Columbia. Practices regarding terminology and definition were determined and the findings are highlighted in this study and compared to prior findings. Most significant was the confirmation of the change from the term mental retardation to intellectual disability while modifications in definitions and therefore diagnostic practices are addressed. Implications are noted with special attention to trends among the states and to considerations with regard to practices within the field.

For more than three decades, professionals in the field of intellectual disability (ID) have witnessed a significant amount of review and scrutiny as related to the use of terminology and the development of definitions, which serve as the basis for diagnostic decisions in this field. Trends in the field have certainly been reflected in the periodic revisions of the definition and terminology manuals promulgated by the American Association on Intellectual and Developmental Disabilities (AAIDD) and by other organizations (e.g., American Psychiatric Association, 2013). As noted by Polloway, Patton, Smith, Antoine, and Lubin (2009), these revisions have served as a basis for re-considering key issues. They typically reflect changes in public policy, professional practice, and/or science and understanding.

A series of research papers have addressed the state of affairs with regard to definition terminology. For example, Denning, Chamberlain, and Polloway (2000) reported the first significant state review since Frankenberger (1984). They reported that 26 states (51.0%) continued to use the term mental retardation; other terms used included mental disability, mental impairment, mentally handicapped, and intellectual disabilities. Denning et al. (2000) found that 35 (68.6%) of all states continued to use the Grossman (1983) definition in either an adapted or verbatim version. They also reported that 13 states indicated that no cut-off score was required, 18 required two standard deviations below the mean, and 13 used a score of 70. For adaptive behavior (AB), 49 states indicated that AB consideration was required, with 14 identifying specific practices. Finally, just 12 of the states indicated a formal age criterion for the developmental period with eight of them noting it as age 18.

Polloway et al. (2009) provided a replication of the Denning et al. (2000) study looking again at terminology, definition, and definitional criteria as did Bergeron, Floyd, and Shands (2008). Both studies reported that 27 states used the term of mental retardation at the time of their study. Polloway et al. (2009) also found that 34 states still used either a verbatim or adapted version of the Grossman (1983) definition. Given the common practice of adopting whatever is used...
in the federal statute for subsequent inclusion in state guidelines, they observed that the Grossman-based definition continued to be relied upon. Polloway et al. (2009) and Bergeron et al. (2008) both reported that the clear majority of states required an IQ cut-off score or range; most commonly, the ceiling level was either IQ 70 or two standard deviations below the mean. In terms of adaptive behavior, Polloway et al. (2009) found that essentially all states required some assessment in this area.

The purpose of the current study was to replicate prior research and determine the current status of terminology and definitions. The study sought to accomplish this goal by evaluating website-based data from departments of education in all 51 states (including the District of Columbia). Specific foci therefore included terminology, definition, and thus eligibility criteria field of intellectual disability. The assumption was that these data from the 51 governmental entities would provide a profile of current educational practices for persons with intellectual disability in the United States.

Method

The research for this study was conducted using a data mining procedure for websites of the 50 respective states and the District of Columbia. The initial review of the sites was conducted January-February 2016 by one of the co-authors of this paper. Subsequently, a second review was conducted in May 2016 by another one of the authors. In addition to providing further currency of the data, the second review also provided additional attention to the numerous states that continue to post old guidelines in addition to new guidelines, hence potentially creating confusion in terms of the recency of data. The senior author then reviewed any instances of non-concurrence in order to resolve the accuracy of the data.

In general, the procedure used Google as the search engine to find the eligibility criteria for intellectual disability on the websites for the 51 departments of education. On these sites, special attention was given to the bottom of the website for ‘last updated’ so that there was assurance that this information was correct and was the most recent. Also, there were some state websites that were not apparently recently revised. In that case, a search was conducted on their ‘revised’ state laws and codes. This process was helpful as many of the states’ codes were revised in 2015.

The focus of the analyses was on the following eight considerations. First, determination was made of the specific term used within definitions to designate the relevant population for this study. Second, an analysis was done to determine the nature of the specific definition used within a given state by comparing it to the various professional definitions used in the field with attention to whether the state definition was used verbatim from the professional organization definition or whether it was similar but adapted. These definitions included the following definitions: Grossman (1983) verbatim (the basis for federal legislation including IDEA for several decades); Grossman (1983) in adapted form; Luckasson et al. (1992) verbatim; Luckasson et al. (1992) adapted; Luckasson et al. (2002) verbatim; Luckasson et al. (2002) adapted; Schalock et al. (2010) verbatim; Schalock et al. (2010) adapted; or other. See Table 1 for the coding sheet for these respective definitional options.

The third issue researched was confirmation of the inclusion of reference to intellectual functioning as a definitional (and diagnostic) component of the definition. Related to that question, the fourth question focused on whether the guidelines specified a cut-off score for intellectual functioning and, if an IQ score was specified, the score or score ranges. The fifth issue then focused on whether the state guidelines required consideration of adaptive skills or adaptive behavior. If so, a related, sixth question then focused on whether specific aspects of adaptive behavior were delineated and/or whether assessment practices were recommended.

The final two issues addressed were whether the guidelines included the concept of a specific developmental period within their definition as part of defining the time when the disability originated and, if guidelines indicated a specific age or range, which was then identified.
Results

Data obtained on each of the respective 50 states and the District of Columbia were entered in detail into a spreadsheet that summarized all of the information reviewed. Then, these extensive data were summarized according to the eight issues identified above; the summary is presented in Table 2.

As noted above, the initial item focused on terminology. Specific attention was given to the specific term that was used within each of the respective states. The data indicate that the following pattern currently exists across the 51 entities. The term, intellectual disability, was reported in a total of 42 state guidelines (82.3%). Mental retardation continued to be reported as being used in three states (5.8%). In addition, three states use the term cognitive disability and one state each cited intellectual impairment, developmental cognitive disability, and cognitive impairment.

The second item focused on the professional definition that had been used as the basis of state eligibility guidelines. As noted above, these definitions were coded by two of the authors of this study and followed the information presented in Appendix A. The Grossman (1983) definition was used by 33 (64.7%) states in an adapted version (GA) while no states used it verbatim (largely because the Grossman definition was based on the term mental retardation). The Luckasson (1992) definition was found to be used in an adapted form by two (3.9%) states (LA1). The Schalock et al. (2000) definition was also used by two states in adapted fashion (SA). Further, 14 states (27.4%) used definitions that were specific to that state and which could not be classified according to the coding system used within this research.

The third and fourth items focused on the “first prong” typically included within the definition for intellectual disability. All 51 entities made reference to deficiencies related to intellectual functioning, with 49 specifically citing the requirement of deficits in intellectual functioning with the other two indicating some variation on that theme (i.e., mental functioning, cognitive functioning).

A total of 41 states (80.4%) also included a recommended cut-off point for this dimension or prong. Sixteen states (31.4%) indicated that the score was 70. An additional 16 states specified the range as below two standard deviations (in four instances qualifying it

---

TABLE 1
State Definitions: Coding System

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. GV: Grossman (1983) verbatim exactly as listed here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Mental retardation refers to significantly sub-average general intellectual functioning existing concurrently with deficits in adaptive behavior and manifested during the developmental period”.</td>
</tr>
<tr>
<td>2. GA: Grossman (1983) adapted and similar but not identical to #1 above</td>
<td></td>
</tr>
<tr>
<td>3. LV1: Luckasson et al. (1992) verbatim: exactly as listed here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mental retardation refers to substantial limitations in present functioning. It is characterized by significantly subaverage intellectual functioning, existing concurrently with related limitations in 2 or more of the following applicable adaptive skill areas: communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and work. Mental retardation manifests before age 18.</td>
</tr>
<tr>
<td>4. LA1: Luckasson et al. (1992) adapted and similar but not identical to #3 above</td>
<td></td>
</tr>
<tr>
<td>5. LV2: Luckasson et al. (2002) verbatim: exactly as listed here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mental retardation is a disability characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills. This disability originates before age 18.</td>
</tr>
<tr>
<td>6. LA2: Luckasson et al. (2002) adapted and similar but not identical to #5 above</td>
<td></td>
</tr>
<tr>
<td>7. SV: Schalock et al. (2010) verbatim: exactly as listed here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intellectual disability is characterized by significant limitations both in intellectual functioning and adaptive behavior as expressed in conceptual, social, and practical skills. This disability originates before age 18.</td>
</tr>
<tr>
<td>8. SA: Schalock et al. (2010) adapted and similar but not identical to #7 above</td>
<td></td>
</tr>
<tr>
<td>9. N/A: dissimilar in form to all of the above eight options</td>
<td></td>
</tr>
</tbody>
</table>
# TABLE 2

## Summary of State Response Data

<table>
<thead>
<tr>
<th>State</th>
<th>Term</th>
<th>Def Type</th>
<th>IF</th>
<th>IQ</th>
<th>AB</th>
<th>AB Specifics</th>
<th>DP</th>
<th>DP age/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Alaska</td>
<td>CI</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Arizona</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Arkansas</td>
<td>MR</td>
<td>GA</td>
<td>yes</td>
<td>70–75</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>California</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Colorado</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2 SD</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Connecticut</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>69</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Delaware</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70 +/−5</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>22</td>
</tr>
<tr>
<td>District of Columbia</td>
<td>ID</td>
<td>LA1</td>
<td>yes</td>
<td>2 SD</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Florida</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>69</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Georgia</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Hawaii</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>2SDs</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>0–8</td>
</tr>
<tr>
<td>Idaho</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Illinois</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Indiana</td>
<td>CD</td>
<td>N/A</td>
<td>yes</td>
<td>2SD</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Iowa</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Kansas</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Kentucky</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Louisiana</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Maine</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Maryland</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>II</td>
<td>N/A</td>
<td>Yes</td>
<td>cog</td>
<td>func</td>
<td>–</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Michigan</td>
<td>MR</td>
<td>LA1</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Minnesota</td>
<td>DCD</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>22</td>
</tr>
<tr>
<td>Mississippi</td>
<td>ID</td>
<td>SA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Missouri</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2SD</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Montana</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Nebraska</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Nevada</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>New Hamp</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>New Jersey</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>60–69</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>New Mexico</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2SD</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>New York</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>North Carolina</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>North Dakota</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>–</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Ohio</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70–75</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70–75</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Oregon</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>MR</td>
<td>GA</td>
<td>yes</td>
<td>2SD</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>South Carolina</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2SDs +/−SEM</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>South Dakota</td>
<td>CD</td>
<td>GA</td>
<td>yes</td>
<td>2SD +/−SEM</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Tennessee</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2SD</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>Before 18</td>
</tr>
<tr>
<td>Texas</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>0–18</td>
</tr>
<tr>
<td>Utah</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>75</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Vermont</td>
<td>ID</td>
<td>N/A</td>
<td>yes</td>
<td>1.5SD</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>Virginia</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
<tr>
<td>Washington</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2SD</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>–</td>
</tr>
<tr>
<td>West Virginia</td>
<td>ID</td>
<td>GA</td>
<td>yes</td>
<td>2SD +/−SEM</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>18</td>
</tr>
</tbody>
</table>
with reference to standard error of measurement), which approximates the score 70 as well. Further, a number of other stipulations were indicated including: below 1.5 standard deviations from the mean (one state), 69 or below (two states), 70–75 (three states), 75 (one state), and 70–75 (two states).

The next two issues addressed concerns about adaptive behavior within the definition. In the current study, all 51 confirmed that consideration of adaptive behavior was required within their definitions. The related issue researched was whether specific aspects of adaptive behavior were delineated and/or whether assessment practices were recommended with 33 states (64.7%) including such information. The information included was determined to be primarily explanatory in terms of the concept of adaptive behavior or further information in detail was provided with guidelines for assessment of adaptive behavior.

The final considerations relate to the inclusion of, and specific stipulation of, a developmental period for the disability. Fifty states cited that the disability must originate in the developmental period. A total of 36 states (70.6%) did designate an age or range for the developmental period including 31 states (60.8%) that indicated 0–18 or through age 18, four states (7.8%) stipulating through age 22, and one state indicating ages 0–8.

Discussion

This review of state guidelines provides a replication of previous research on definitional patterns as reported by Frankenberger (1984), Denning et al. (2000), Bergeron et al. (2008) and Polloway et al. (2009), among others. The sections below highlight each of the points that were researched within this analysis, placing them in the context of trends within the field.

Terminology

The data reported herein confirm the significant changes that have taken place over the last ten years in terminology within the field. A decade ago, Smith (2006) noted “the 2002 AAMR manual authors acknowledged the problems with the term mental retardation but concluded nevertheless that there was no acceptable alternative term, despite its acknowledged shortcomings” (p. 58). In fact, this continued reliance on this term was underscored by limited change between the reports by Polloway et al. (2009; 27 states) vs. the 26 states that Denning et al. (2000) reported using MR. In the current study, only three states retained the MR term with the clear preference now being for intellectual disability (42 states). It is important to note that this shift to ID (from MR) accelerated with the October 5, 2010, signing of Rosa’s Law, a dramatic federal manifestation of the movement.
to eliminate the “R word” (Polloway et al., 2017). The Congressional Record provided a synopsis of the rationale for change (adapted below):

The term “mental retardation,” . . . to describe individuals with intellectual disabilities <is> anachronistic, needlessly insensitive and stigmatizing, and clinically outdated. Terms have gone through a steady evolution over the past two centuries, each iteration describing those living with the condition in a pejorative way. . . . “Imbecile,” “moron,” “idiot,” and “feeble-minded” are all terms which have been used . . . . Each of these terms focused on perceived deficiencies to describe such individuals. The most recent term—“mental retardation”—was used to characterize those with cognitive disabilities as having general diminished capacities for cognitive functioning. . . . Within the past 30 years, . . . “mental retardation” and “mentally retarded” have also developed into colloquial slurs and pejorative phrases used to demean and insult both persons with and without disabilities. Congress has recognized that these negative attributions towards people with disabilities should not be tolerated . . . it is also essential that we ensure these individuals are provided the respect they deserve as part of our American family. Thus it is important to revise our terminology in Federal statutes, as appropriate, to further and support the equality of all individuals, without regard to disability (US Senate 111–244, August 3, 2010).

The increased use of the ID term can be further noted in organizational names (e.g. American Association on Intellectual and Developmental Disabilities), journals (e.g., American Journal of Intellectual and Developmental Disabilities, Intellectual and Developmental Disabilities), professional text titles, and definitions. Polloway, Patton, and Smith (2015) noted that ID has now been almost universally embraced as the preferred term in the field.

Definition

Two recent, nationally recognized sources for the definition, and for diagnostic criteria, for intellectual disability were promulgated by the American Association on Intellectual and Developmental Disabilities (AAIDD; Schalock et al., 2010) and the Diagnostic and Statistical Manual of Mental Disorders of the American Psychiatric Association (DSM-V; APA, 2013; Tasse, 2015). However, these definitions have not had an impact on state practices in education to date; the Schalock et al. definition was found in use in two states (in adapted form) and DSM in no state definitions.

The current study found that clearly the most common source remains the Grossman (1983) adaptation, which appears in IDEA 2004. Naturally, it is common that states mirror the statutory language provided in the federal law for state education statute. In the Polloway et al. (2009) study, 34 states relied on this source, which was somewhat less than the 44 reported by Denning et al. (2000). In the current study, 33 states-virtually the same number as in 2009- continued to rely on this federal definition that is now approximately four decades old.

Intellectual Assessment

In this domain, all states indicated the core element of intellectual functioning included within definitional considerations. Further, 41 states indicated a cutoff score or score range as a basis for confirmation of deficits and intellectual functioning. This figure is comparable to the finding of Polloway et al. (2009) that 35 states indicated requiring a cut-off score or range. Bergeron et al. (2008) and Polloway et al. (2009) reported that the clear majority of states required a cut-off score or range; most commonly, the ceiling level was either approximately IQ 70 or two standard deviations below the mean. Again, in the current study, most commonly found was that the the respective states used a score of 70 or approximately 70 or the relative equivalent of two standard deviations below the mean, as the ceiling level for intellectual disability. These data also echo those reported by Denning et al. (2000) that 31 states used a score of 70 or 2 SDs as the cut-off score.

While precise cut-off scores have varied by definition and by state guidelines (Polloway et al., 2010), nevertheless the use of a range of scores reflecting, for example, error in measurement is the professionally accepted position and can be compared to the so-called “bright line” interpretations of intellectual disability in which some states used a strict IQ ceiling within criminal justice arena, most commonly 70. In Hall v Florida (2014), the Supreme Court rejected the use of such a
bright line cutoff (Polloway et al., 2015). In the words of Justice Kennedy in that case, “intellectual disability is a condition, not a number”.

**Adaptive Behavior Assessment**

The second prong of ID is adaptive behavior (AB). Within this domain, the contemporary concept of adaptive behavior is typically based on a tripartite competence model, which gave rise to the identification of the three domains including conceptual, practical, and social adaptive behavior (Greenspan, 2015). Contemporary practice requires that an individual show significant deficits in overall AB or within conceptual, social, or practical adaptive skills. However, this concept of adaptive behavior and this level of discrimination implementing the concept have not commonly been reflected in definitional efforts across the 51 states as reported herein. Nevertheless, in the current study, all states did require consideration of adaptive behavior in defining intellectual disability, which was also reported by Polloway et al. (2009) and by Denning et al. (2000). Further, a total of 33 states or 64.7% did provide additional guidance for implementation of the concept of adaptive behavior within diagnostic decisions intellectual disability. These data do reflect an increase from the 23 states reported by Polloway et al. (2009) and the 14 states as reported by Denning et al. (2000). Some variance in coding this information does require caution in concluding that there is a significant trend towards increased specificity in terms of adaptive behavior assessment.

**Developmental Period**

The third common prong of the definition of intellectual disability relates to age of origination or onset, which focuses attention therefore on the developmental aspect of this disability. The two most contemporary definitions continue this commitment that has been present over perhaps five decades in the consideration of intellectual disability. The AAIDD definition (Schalock et al., 2010) indicates that intellectual disability be manifested during the period of 0–18 years. DSM-V (APA, 2013) also adheres to the developmental period defined as ages 0–18.

In the current study, virtually all states (98.0%) affirmed the fact that intellectual disability is a developmental phenomenon and that disability should be manifested during the developmental period. Further, consistent with the definitions noted above, the predominant age range for the developmental period was confirmed as 0–18 years (60.8%) with four states extending this period through age 22. Of note is the fact that in the current study only 29.4% of all states did not specify relevant ages for the developmental period, which represent a significant decrease from the findings reported by Polloway et al. (2009) that the majority of states (33 or 64.7%) did not specify an age or range and the 76.5% reported by Denning et al. (2000).

**Limitations**

Several limitations should be acknowledged in the study. First, the information used in this research was based on data obtained from Internet sources for the respective 51 entities. There was not a confirmation received from special education directors within the respective states concerning the accuracy of their posted data. Further, the data are necessarily time-limited because there was no way to determine whether individual states may have been in the process of making changes in the near future. While acknowledging the relevance of these limitations, the current study reported herein does nevertheless contribute to greater understanding of educational practices in the field of intellectual disability on a national level.

**Implications**

There are several implications that derive from the current study. First, the data confirm the significant trend away from use of the term mental retardation across almost all states and the significant increase in reliance on intellectual disability, which is now also reflected in federal regulations. Second, the study highlights the continued reliance on the primary prongs of the definition intellectual disability including considerations of intellec-
tual functioning, adaptive behavior, and manifestation during the developmental period.

Third, the data presented illustrate the fact that the federal definition used for intellectual disability is essentially one that has been adopted by the majority of states even though it has been largely unchanged (beyond the term itself in the definition) for over 30 years. It would seem most appropriate to consider contemporary views of intellectual disability, such as reflected in the current Schalock et al. (2010) definition as a basis for updating both federal and state definitions.

References


Received: 28 July 2016
Initial Acceptance: 27 September 2016
Final Acceptance: 6 November 2016
Differentiating Instruction in the Inclusive Classroom

STRATEGIES FOR SUCCESS

BARBARA GARTIN
NIKKI MURDICK
DARLENE PERNER
MARcia IMBEAU
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. Three copies of the manuscript along with a transmittal letter should be sent 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.
CEC’s Division on Autism and Developmental Disabilities (DADD) is hosting the 19th International Conference, which will focus on research-informed practice. DADD is pleased to extend an invitation to join us in Clearwater Beach, Florida, January 17 - 19, 2018, for a stellar professional learning opportunity! The program will feature more than 200 lecture and poster presentations in the following topical areas:

- Autism Spectrum Disorder
- Assistive & Adaptive Technology
- Early Childhood
- Intellectual Disability
- Mental Health
- Multiple Disabilities
- Paraprofessionals
- Employment
- Post-Secondary
- Transition

For further information, please contact:
Cindy Perras, Conference Co-ordinator
cindy.perras@gmail.com
www.daddcec.org