Education and Training in Autism and Developmental Disabilities

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

Volume 50 Number 2

June 2015
Education and Training in Autism and Developmental Disabilities

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

Editorial Assistant: Jessica Holloway-Libell
Arizona State University
Mary Lou Fulton Teachers College

Consulting Editors
Martin Agran
Kevin Ayres
Devender Banda
Laura Bassette
Emily Bouch
Amanda Boutot
Michael P. Brady
Stacy Carter
David Cihak
Ginevra Courtade
Samuel A. DiGangi
Teresa Doughby
Jennifer Ganz
Juliet Hart Barnett
Carolyn Hughes
Bree Jimenez
Russ Lang
Justin Leaf
Wendy Machalicek
Michael Mayton
John McDonnell
Hedda Meadan-Kaplansky
Linda Mechling
Pam Mims
Wendy Oakes
Karrie Shogren
Scott Sparks
Fred Spooner
Robert Stodden
Keith Storey
William Therrien
Matt Tincani
Jason Travers
Tony Van Laarhoven
Elizabeth West
John Wheeler
Mark Wolery
Kimberly W. Fisher
Leah Wood
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-psych, 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

<table>
<thead>
<tr>
<th>Officers</th>
<th>Members</th>
<th>Executive Director</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past President Amanda Boutot</td>
<td>Debra Cote</td>
<td>Teresa Doughty</td>
</tr>
<tr>
<td>President Diane Zager</td>
<td>Richard Gargiulo</td>
<td>Publications Chair</td>
</tr>
<tr>
<td>President-Elect David Cihak</td>
<td>Beth Kannagah</td>
<td>Michael Wehmeyer</td>
</tr>
<tr>
<td>Vice President Elizabeth West</td>
<td>Jenny Root (Student Representative)</td>
<td>Communications Chair</td>
</tr>
<tr>
<td>Secretary Dagny Fidler</td>
<td>Lynn Stansberry-Brunahan</td>
<td>Emily C. Bouch</td>
</tr>
<tr>
<td>Treasurer Gardner Umbarger</td>
<td>Angie Stone-MacDonald</td>
<td>Conference Coordinator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cindy Perras</td>
</tr>
</tbody>
</table>

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. $60.00 per year; Canada, PUAS, and all other countries $64.00; Institutions—U.S. $195.00 per year; Canada, PUAS, and all other countries $199.50; single copy price is $30.00. 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.
What is the Proof? A Methodological Review of Studies That Have Utilized Social Stories
JUSTIN B. LEAF, MISTY L. OPPENHEIM-LEAF, RONALD B. LEAF, MITCHELL TAUBMAN, JOHN McEACHIN, TRACEE PARKER, ANDREA B. WAKS, and TOBY MOUNTJOY

Compensation Age Theory: Effect of Chronological Age on Individuals with Intellectual Disability
HEFZIBA LIFSHITZ-VAHAV

Incorporating Functional Digital Literacy Skills as Part of the Curriculum for High School Students with Intellectual Disability
DAVID F. CIHAK, RACHEL WRIGHT, CATE C. SMITH, DON McMAHON, and KELLY KRAISS

Review of Evidence-based Mathematics Interventions for Students with Autism Spectrum Disorders
JULIET E. HART BARNETT and SHANNON CLEARY

Is There Really a Difference? Distinguishing Mild Intellectual Disability from Similar Disability Categories
EMILY C. BOUCK and RAJIV SATSANGI

Utility of Formal Preference Assessments for Individuals Diagnosed with Autism Spectrum Disorder
JUSTIN B. LEAF, RONALD LEAF, ADITT ALACALAY, JEREMY A. LEAF, DANIEL RAVID, STEPHANIE DALE, ALYNE KUYUMJIAN, KATHLEEN TSUJI, MITCHELL TAUBMAN, JOHN McEACHIN, and MISTY OPPENHEIM-LEAF

Use of Say-Do Correspondence Training to Increase Generalization of Social Interaction Skills at Recess for Children with Autism Spectrum Disorder
NANCY ROSENBERG, MARISSA CONGDON, ILENE SCHWARTZ, and DEBRA KAMPS

An Investigation of the Effects of CRA Instruction and Students with Autism Spectrum Disorder
SHAUNITA STROIZER, VANESSA HINTON, MARGARET FLORES, and LATONYA TERRY

Effect of Instruction with the Self-Determined Learning Model of Instruction on Students with Disabilities: A Meta-Analysis
SUK-HYANG LEE and MICHAEL L. WEHMEYER

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

The Division on Autism and Developmental Disabilities retains literary property rights on copyrighted articles. Up to 100 copies of the articles in this journal may be reproduced for nonprofit distribution without permission from the publisher. All other forms of reproduction require permission from the publisher.
Manuscripts Accepted for Future Publication in Education and Training in Autism and Developmental Disabilities

September 2015

Use of a proximity sensor switch for “hands free” operation of computer-based video prompting by young adults with moderate intellectual disability. Alexandria N. Ivey, Linda C. Mechling, and Galen P. Spencer, University of North Carolina Wilmington, Department of Early Childhood & Special Education, 601 South College Road, Wilmington, NC 28403-5940.


Using science inquiry methods to promote self-determination and problem-solving skills for students with moderate intellectual disability. Bridget Miller, Teresa Doughty, and Gerald Krockover, Department of Teaching and Instruction, University of South Carolina, 820 Main Street, Columbia, SC 29208.

Increasing social engagement in an inclusive environment. Rebecca Hartzell, Candace Gann, Carl Liaupsin, and Sarah Clem, 3559 West Songbush Place, Marana, AZ 85658.

Review of sibling interventions with children with autism. Devender Banda, College of Education, PO Box 41071, Texas Tech University, Lubbock, TX 79409.

Comparison of three video perspectives when using video prompting by students with moderate intellectual disability. Galen P. Spencer, Linda C. Mechling, and Alexandria N. Ivey, University of North Carolina Wilmington, 601 S. College Road, Wilmington, NC 28403.

Animal assisted therapy for children with autism spectrum disorders: A systematic review. Tonya N. Davis, Rachel Scalzo, Erin Butler, Megan Stauffer, Yara N. Farah, Scott Perez, Kristen Mainor, Cathryn Clark, Stacy Miller, Alicia Kobylecky, and Laura Coviello, Baylor University, One Bear Place #97301, Waco, TX 76798-7301.

Effects of a graphic organizer training package on the persuasive writing of middle school students with autism. Anne E. Bishop, Mary Sawyer, Sheila R. Alber-Morgan, and Melissa Boggs, The Ohio State University, A356 PAES Building, 305 W. 17th Avenue, Columbus, OH 43210.

Using an animated cartoon hero in video instruction to improve bathroom-related skills of a student with autism spectrum disorder. Yoshihisa Ohtake, Ayaka Takahashi, and Kentaro Watanabe, Division of Developmental Studies and Support, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama-shi, Okayama, 700-8530, JAPAN.

Address is supplied for author in boldface type.
What is the Proof? A Methodological Review of Studies That Have Utilized Social Stories

Justin B. Leaf, Misty L. Oppenheim-Leaf, Ronald B. Leaf, Mitchell Taubman, John McEachin, Tracee Parker, Andrea B. Waks, and Toby Mountjoy
Autism Partnership Foundation

Abstract: Social stories are a commonly empirically evaluated and implemented procedure to increase pro-social behaviors and decrease aberrant behaviors for individuals diagnosed with an autism spectrum disorder. Despite their widespread use there have been questions raised to the soundness of the research methodology and the results which have been demonstrated within these research studies. This paper is a methodological review of 41 studies that evaluated social stories for individuals diagnosed with autism. We classified each study as one that utilized either a case study design, a reversal design, or a multiple baseline design. After classification we evaluated each study across multiple methodological dimensions and used this analysis to determine if a study showed either a clear demonstration, partial demonstration, or if there was no clear demonstration that the social story was responsible for behavior change. Results of this analysis indicated that the majority of studies either showed only a partial demonstration or no clear demonstration that the social story procedure was responsible for the behavior change. Based upon this analysis recommendations for clinicians and future researchers are discussed.

In 1993, Gray and Garand introduced social stories as a strategy that could be used to increase positive behaviors and reduce aberrant behaviors for individuals diagnosed with autism spectrum disorder (ASD). Social stories are a systematic form of intervention in which a teacher writes a brief text that describes a targeted behavior to be displayed by the student, when the student should display the desired behavior, why the student should display the targeted behavior, and how displaying the behavior will affect others in his or her environment.

In describing social stories, Gray and Garand specified several important guidelines. First, social stories should only be implemented with children and adolescents who fall in the “trainable mentally impaired range or higher who possess basic language skills” (Gray & Garand, 1993, p. 103). Second, each social story should be individualized. Third, social stories should only address one behavior at a time. Fourth, Gray and Garand suggested that, when reading a social story, the teacher and the participant should sit side by side. Fifth, social stories should use only positive language. Finally, they suggested specific types of sentences that should be used within a social story (Gray & Garand, 1993; Gray, 1994, 2004). Each social story should consist of at least four sentence types (i.e., descriptive, perspective, affirmative, and directive). Gray (2000, 2002, 2003) later added additional sentence types, which included: cooperative sentences, control sentences, and partial sentences. Gray (1995) later recommended that for every directive sentence there should be a total of two to four of the other sentence types.

Since the original article (i.e., Gray & Garand, 1993), there have been numerous articles published that have evaluated social stories on teaching social behavior (e.g., Adams, Gouvouis, VanLue, & Waldron, 2004; Barry & Burlew, 2004; Crozier & Tincani, 2007), teaching behaviors that are not inherently social (e.g., sitting) (e.g., Bledsoe, Myles, & Simpson, 2003; Hagiwara & Myles, 1999), and on decreasing aberrant behavior(s) (e.g., Agosta, Graetz, Mastroplieria, & Scruggs, 2004 Chan & O’Reily, 2008; Crozier & Tincani, 2005;
Lorimer, Simpson, Myles, & Ganz, 2002; Kuttler, Myles, & Carlson, 1998). Additionally, there have been several curriculum books written on how to write and effectively implement social stories (Gray, 2000, 2002, 2003). Finally, several reviews have been published examining the effectiveness and efficiency of social stories (e.g., Kokina, & Kern, 2010; Sanosti, Powell-Smith, & Kincaid, 2004).

One of the first reviews on social stories was written by Sansosti et al, (2004), which evaluated studies that implemented social stories using single subject methodology. Sansosti and colleagues determined that the effects of social stories were limited in the articles they reviewed, due to possible confounding variables, weak methodological control, and weak treatment effects. Ali and Frederickson (2006) also acknowledged the limitations of the social story research, but stated that “the approach has promise and warrants further research” (Ali & Frederickson, 2006, p. 372).

In a more recent review, Kokina and Kern (2010) evaluated 18 different studies of social stories on a variety of dimensions (e.g., age, diagnosis, number of sessions, story format, use of comprehension questions). Similar to previous reviews, the authors concluded that research on social stories has questionable results in terms of effectiveness, and that the social story research has methodological flaws, which further limits the results.

Despite the lack of empirical evidence demonstrating efficacy, social stories are implemented with high frequency by teachers, parents and clinicians to individuals diagnosed with ASD (Reynhout & Carter, 2009). Reynhout and Carter (2009) surveyed 45 teachers regarding the perceived efficacy of social stories. The authors used a questionnaire with a 5-point Likert scale to examine various questions about social stories. The results from this survey revealed that 93.4% of the teachers surveyed considered social stories to be an effective intervention (48.9% agreed and 48.9% strongly agreed). Furthermore, the National Standards Report Project (The National Autism Center, 2009) for evidence based practices stated that story based intervention packages is an established procedure. Thus, it appears, that despite methodological flaws, social stories remains a widely used and endorsed procedure.

When evaluating research, it is imperative not only to look at the effects of the intervention but to evaluate if the researchers implemented the methodology appropriately. For example, if the researchers implemented the intervention condition at the wrong time (e.g., the data in baseline trending the desired way of intervention), it would not be clear if the change in behavior was due to the intervention or if the change in behavior was going to occur naturally. The purpose of this paper is to evaluate the correct use of single subject designs in research that utilized social stories for individuals diagnosed with autism. To do this, we evaluated research published in peer reviewed journals from the years 1993 (i.e., the year of the first publication on social stories) to 2012. This paper will specifically look at case study designs, reversal designs, and multiple baseline designs and evaluate the experimental rigor that was utilized. We will also utilize visual inspection to determine whether a clear behavior change was shown across the different studies. Based on the soundness of the research methodology and the extent of behavior change that was demonstrated, we will determine the documented effectiveness of the social story procedure for changing the participant(s) behavior(s). Finally, we will make recommendations to clinicians and researchers.

Search Procedure and Inclusion Criterion

Two researchers independently conducted multiple searches to identify studies that implemented social stories with individuals diagnosed with ASD. First, we conducted electronic searches on ERIC and PsycINFO using several combinations of keywords. Second, we conducted a visual search of the following journals: Journal of Applied Behavior Analysis, Journal of Autism and Developmental Disabilities, Research in Autism Spectrum Disorders, Focus on Autism and Other Developmental Disabilities, and Autism. We reviewed these journals as they commonly publish interventions for individuals diagnosed with autism. Third, we examined any references that were cited directly in the articles that were retrieved.

To be included in this review each research study had to meet four inclusion criteria. First, the article had to include at least one partici-
pant diagnosed with ASD. Second, the study had to utilize single subject methodology. Third, any comparative study was not included. Finally, all research studies had to be published or made available online in a peer reviewed journal from January 1993 to December 2012. After the two independent reviews, the two researchers identified a combined 41 studies that met this criterion. Next, the two researchers independently assigned each of the 41 studies to case-study design, reversal design, or multiple baseline design. The two researchers demonstrated 100% agreement regarding the categorical assignment of each research article.

**Designs**

*Case Study Design*

Nine studies were classified as a case study design. Case study designs are the most basic of all single subject designs. The most commonly implemented case study design is the AB case study design, which reports baseline data prior to intervention. There also exists a B (intervention) only case study design, but it provides no way of judging the impact of an intervention as no information is provided regarding baseline performance. The AB case study design begins with an initial baseline period (A) when the researcher measures the behavior of the participant prior to the intervention. Once the participant demonstrates consistent responding or responding trending in the opposite direction to that of the desired treatment effect, the researcher implements the treatment condition (B). Intervention is considered to be “successful” if the learner’s behavior rapidly and significantly changes during the intervention condition.

Although the case study design is considered part of the single subject methodology, it is a weak design (Bailey & Burch, 2002). This is because, without replication either within or across participants, the case study design does not rule out many confounding variables (e.g., history, incidental occurrences, maturation, etc.) and it is difficult to determine if the intervention was truly responsible for the behavior change or if the behavior change was due to some other variable. Researchers can maximize the believability of this design in several ways. First, they can conduct a longer baseline period. Second, the baseline data must be stable or trending in the opposite direction of the desired treatment effect prior to intervention. Third, there should be little to no overlap between the baseline data and intervention data. Finally, the participant’s behavior change should occur rapidly once the intervention condition begins (Bailey & Burch, 2002).

*Reversal Designs*

Thirteen studies were classified as utilizing a reversal design (e.g., ABA, ABAB design, ABABAC design). The reversal design is the most robust of all of the single subject designs. Essentially, the design demonstrates that the targeted behavior can be “turned on” or “turned off” at the researcher’s will. The design typically starts with an initial baseline period (A) where the researcher measures the behavior of the participant prior to the intervention. Once the participant demonstrates consistent responding or responding trending in the opposite direction to that of the desired treatment effect, the researcher implements the intervention condition (B). Following implementation of the intervention, once the participant responds at a stable rate or displays the behavior at a rate trending in the desired direction, the researcher switches back to the baseline condition. The researcher continues to reverse the conditions to demonstrate functional control.

In order to show functional control the researcher must adhere to certain guidelines. First, in all baseline conditions the behavior must be stable or trending in the opposite direction of the desired treatment effect before switching conditions. Second, in all intervention conditions the behavior needs to be stable, trending in the desired direction, and at higher levels than during the baseline condition, prior to switching back to the baseline condition. Third, researchers should only implement this design with behaviors that can easily be reversed and for behaviors that will not cause injury to the participant or to others.

*Multiple Baseline Designs*

Nineteen of the studies were classified as utilizing a multiple baseline design or a variation
of the multiple baseline design (e.g., multiple probe design). The multiple baseline design is well suited when the target behavior is nonreversible or when it would be unethical to reverse the targeted behavior. A multiple baseline design is used when a researcher wants to measure an intervention across two or more participants, two or more behaviors for the same participant, or two or more settings (e.g., home and school). When the researcher implements this design they start by collecting baseline measures for each of the participants, behaviors, or settings. The researcher then implements intervention in a stepwise fashion across the multiple participants, behaviors, or settings. The researcher does not implement intervention for a new participant, behavior, or setting until there has been a significant change of behavior for the previous participant, for the previous behavior, or in the previous setting. Additionally, the researcher should not implement each intervention until the baseline behavior is stable or trending in the correct direction. This way, each baseline demonstrates that the intervention is responsible for the behavior change; more baselines (e.g., more participants) allows for the opportunity to replicate the treatment effects.

Measures across Research Designs

Measures for Case Study Design

We evaluated the nine studies that utilized the case study design across six main variables. First, we identified how many sessions were conducted in the baseline condition for each skill targeted in each of the studies. Second, we assessed whether the baseline data was stable or trending in the correct direction prior to switching to the intervention condition. The definition of stability or trending in the correct direction consisted of four components: (a) the participant displaying the same rate of behavior for the final two baseline sessions prior to switching to the intervention condition; (b) the participant displaying a rate of behavior that is trending in the opposite direction of the desired treatment effect for the final two baseline sessions prior to switching to the intervention condition; (c) the participant displaying the same rate of behavior or behavior at a rate trending in the opposite direction of the desired treatment effect for three out of the last four baseline sessions prior to switching to the intervention condition, without the last data point trending in the wrong direction prior to the switch; and (d) if two or more dependent variables were being measured, all measures had to meet the criterion stated above unless the researcher specifically stated that one of the dependent variables was the main dependent variable.

Third, we evaluated whether the treatment effect occurred immediately. This was defined as the behavior changing in the desired direction for two of the first three intervention sessions, or the third data point being higher (for increasing behaviors) or lower (for decreasing behaviors) than all data points in the baseline condition. Fourth, we compared the number of intervention data points that overlapped with the baseline data or that was trending in the wrong direction of the desired treatment effect. Fifth, we evaluated if the researchers used objective or subjective data. Finally, we evaluated whether the social story procedure was combined with any other procedure (e.g., video modeling, prompting, scripts) at any point within the study.

Measures for Reversal Design

We evaluated the 13 studies that utilized reversal design across five main variables. First, we calculated the percentage of all baseline conditions where the participant’s behavior(s) were stable or trending in the opposite direction of the desired treatment effect. The same definition was utilized as for the case-study design. We calculated the number of baseline conditions where stability or trending in the correct direction was shown and divided by the total number of baseline conditions where an intervention condition directly followed the baseline condition.

Second, we calculated the percentage of all intervention conditions where the participant’s behavior was stable or trending in the desired direction prior to reintroduction of a baseline condition. This definition consisted of three components: (a) the last two data points, prior to switching back to the baseline condition, was stable or trending in the direction of the desired treatment condition and higher than 85 % of all of the baseline data
points or (b) three of the last four data points, prior to switching back to the baseline condition, was higher than 85% of all baseline data points and the last data point was not trending in the opposite direction of the desired treatment effect; and (c) if two or more dependent variables were being measured, all measures had to meet the criterion stated above unless the researcher specifically stated that one of the dependent variables was the main dependent variable. We calculated the number of intervention conditions where stability or trending in the correct direction was shown and divided by the total number of intervention conditions where a baseline condition directly followed the intervention.

Third, we utilized visual inspection to determine the percentage of intervention conditions where there was a clear behavior change from the baseline conditions. This definition consisted of three components: (a) for positive behaviors, 75% of all intervention data points demonstrated an increase from all baseline data points; or (b) for aberrant behaviors, 75% of all intervention data points demonstrated a decrease from baseline behaviors; or (c) a clear level change in the behavior could be observed through visual inspection. We calculated the number of intervention conditions where a change in behavior was demonstrated divided by the total number of intervention conditions.

Level of Demonstration (Analysis of Measures)

We used the measures described above to create three levels of demonstration for studies in each research design category. This analysis was used to identify whether or not the researchers implemented the single subject methodology with enough evidence to convince the reader that the social story procedure was responsible for the behavior change. The three levels of demonstration were: 1) convincing demonstration; 2) partial demonstration; 3) no convincing evidence that the social story procedure was responsible for changing the targeted behavior. Across all studies, if the researcher(s) did not report a measure (e.g., not report baseline data) then that measure was considered as not occurring. In addition, a study had to meet all criteria of a demonstration level (see Table 1, 2, and 3) in order to be characterized as that level of demonstration. If a study met criteria from different demonstration levels, that study was categorized as the lowest level of demonstration for which it met all criteria. All criteria were considered necessary in order to demonstrate that the social story procedure was responsible for behavior change; thus, not meet-
ing any single criterion was considered a flaw of the study’s demonstration that the social story procedure was responsible for the behavior change. Table 1 displays the different scoring criterion and levels of demonstration for studies where a case study design was utilized. Table 2 displays the different scoring criterion and levels of demonstration for studies where a reversal design was utilized. Table 3 displays the different scoring criterion and levels of demonstration for studies where a multiple baseline design was utilized.

Results
Table 4 provides the overall results on the level of demonstration for studies across the three designs evaluated. Table 5, 6, and 7 provide information about how each of the studies scored across the measures assessed and the level of demonstration each research study was classified as. Within each table the rationale for why a study was classified as a given level is highlighted within a cell (category). If any study does not have any highlights in any cell it means that study was determined to have a convincing level of demonstration.

Levels of Demonstration

<table>
<thead>
<tr>
<th>Level of Demonstration</th>
<th>Type of Data</th>
<th>Length of Baseline</th>
<th>Baseline Trending</th>
<th>Effect Immediate</th>
<th>Overlapping Data</th>
<th>Combined with other procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convincing Evidence</td>
<td>Objective</td>
<td>3 or more sessions of baseline</td>
<td>Stable or trending in correct direction</td>
<td>Behavior change demonstrated within 3 sessions</td>
<td>20-40% overlapping data between baseline and intervention</td>
<td>Not combined with other procedures</td>
</tr>
<tr>
<td>Partial Evidence</td>
<td>Objective</td>
<td>1 or 2 sessions of baseline</td>
<td>Stable or trending in correct direction</td>
<td>Behavior change demonstrated within 3 sessions</td>
<td>40-61% overlapping data between baseline and intervention</td>
<td>Combined with other procedures</td>
</tr>
<tr>
<td>No Convincing Evidence</td>
<td>Subjective</td>
<td>0 sessions of baseline or baseline not reported</td>
<td>No stability or not trending in the correct direction</td>
<td>Behavior change occurring after 3 sessions</td>
<td>100 to 100% overlapping data between baseline and intervention</td>
<td>Combined with other procedures</td>
</tr>
</tbody>
</table>

TABLE 1
Measures and Demonstration Levels for Case Study Designs

In 2008, Chan and O'Reilly published the second study that showed a clear demonstration of the social story's effectiveness. In this study, the authors utilized a multiple baseline design across three participants to show that the social story was effective in changing the behavior of participants diagnosed with autism. In this study, the authors demonstrated that social stories could be used to teach appropriate social engagement while decreasing inappropriate social engagement or no social engagement.

In 2009, Delano and colleagues published the first study that showed a clear demonstration of the social story's effectiveness. In this study, the authors utilized a multiple baseline design across three participants to show that the social story was effective in changing the behavior of participants diagnosed with autism. In this study, the authors demonstrated that social stories could be used to teach appropriate social engagement while decreasing inappropriate social engagement or no social engagement.

In 2010, Delano and colleagues published the second study that showed a clear demonstration of the social story's effectiveness. In this study, the authors utilized a multiple baseline design across three participants to show that the social story was effective in changing the behavior of participants diagnosed with autism. In this study, the authors demonstrated that social stories could be used to teach appropriate social engagement while decreasing inappropriate social engagement or no social engagement.
tion that social stories were effective in changing behaviors. The social story did include role-plays after the story was read to the students. The researchers utilized a multiple probe design across behaviors (e.g., raising hand, social interaction, and vocalizations) to change two students’ behaviors.

In 2011, Richter and Test published the final study that showed a clear demonstration that social stories were effective in changing participants’ behaviors. In this study the authors utilized a multimedia social story to teach adults how to answer questions on a 16 question quiz. The researchers utilized a multiple probe design across participants and results showed a clear increase in the participants’ ability to answer questions correctly; however, there was no measure of participants’ behavior other than answering questions on the quiz.

**Partial Demonstration.** Out of the 41 studies evaluated, 17 studies (41.4%) displayed a partial level of demonstration that the social story was responsible for the participant changing his or her behavior. None of the studies that utilized a case study design were able to show partial demonstration. Seven of the 13 studies (53.8%) that utilized a reversal design were able to show partial demonstration and 10 out of the 19 studies (52.6%) that utilized a multiple baseline design were able to show partial demonstration.

There were seven studies that used a reversal design that were classified as partial demonstration. Four of these studies implemented multiple procedures (i.e., music, prompts, and reminders) along with the social story procedure. Four of the studies, at least once, introduced the intervention when the participant’s baseline behavior was not stable or was trending in the wrong direction. Three of these studies, at least once, reversed back to the baseline condition when the intervention data was not stable or was trending in the

**TABLE 2**

<table>
<thead>
<tr>
<th>Level of Demonstration</th>
<th>Type of Data</th>
<th>Baseline conditions stable or trending correctly prior to intervention</th>
<th>Intervention conditions stable and higher than baseline or trending the correct way prior to reversal</th>
<th>Clear behavior change during intervention conditions</th>
<th>Combined with other procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convincing Evidence</td>
<td>Objective</td>
<td>100% of all baseline conditions</td>
<td>100% of all intervention conditions</td>
<td>100 to 80% of all intervention conditions</td>
<td>Not combined with other procedures</td>
</tr>
<tr>
<td>Partial Evidence</td>
<td>Objective</td>
<td>99 to 50% of all baseline conditions</td>
<td>99 to 50% of all intervention conditions</td>
<td>79 to 50% of all intervention conditions</td>
<td>Combined with other procedures</td>
</tr>
<tr>
<td>No Convincing Evidence</td>
<td>Subjective</td>
<td>49 to 0% of all baseline conditions</td>
<td>49 to 0% of all intervention conditions</td>
<td>49 to 0% of all intervention conditions</td>
<td>Combined with other procedures</td>
</tr>
</tbody>
</table>

**TABLE 3**

<table>
<thead>
<tr>
<th>Level of Demonstration</th>
<th>Type of Data</th>
<th>Baseline conditions stable or trending in the correct direction prior to intervention</th>
<th>Appropriate staggering of interventions</th>
<th>Clear behavior change in all intervention conditions</th>
<th>Combined with other procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convincing Evidence</td>
<td>Objective</td>
<td>100% of all baseline conditions</td>
<td>100 to 75% correctly staggering intervention</td>
<td>100 to 80% of all intervention conditions</td>
<td>Not combined with other procedures</td>
</tr>
<tr>
<td>Partial Evidence</td>
<td>Objective</td>
<td>99 to 66% of all baseline conditions</td>
<td>74 to 50% correctly staggering intervention</td>
<td>79 to 50% of all intervention conditions</td>
<td>Combined with other procedures</td>
</tr>
<tr>
<td>No Convincing Evidence</td>
<td>Subjective</td>
<td>65 to 0% of all baseline conditions</td>
<td>49 to 0% correctly staggering intervention</td>
<td>49 to 0% of all intervention conditions</td>
<td>Combined with other procedures</td>
</tr>
</tbody>
</table>
wrong direction. Finally, three of these studies did not show a clear change in the participants’ behaviors.

There were 10 studies that used a multiple baseline design that were classified as partial demonstration. Four studies implemented multiple procedures (i.e., video modeling, prompts, video feedback) along with the social story procedure; however, Thiemann and colleagues (2001) was the only study placed in the partial demonstration category for this reason alone. Two studies using a multiple baseline design, at least once, introduced the intervention at an inappropriate time; three did not stagger interventions correctly; and five did not demonstrate improvement in participant behavior.

No Clear Demonstration. Out of the 41 studies evaluated, 21 (51.2%) studies showed no convincing evidence that social stories were responsible for behavior change. Nine of the nine studies that used a case study design were unable to show a clear demonstration and failed to meet multiple criteria to be considered as partial demonstration. Four of the studies utilized additional procedures other than social stories.

There were six studies that used a reversal design that were classified as no convincing evidence that the social story was responsible for the behavior change. Three of these studies were placed in this level of demonstration due to their lacking of behavioral stability or behavior trending in the correct direction during baseline. Three studies were placed in this level of demonstration due to no behavior change demonstrated and one study was placed in this level due to changing from intervention to baseline conditions without behavior stability or trending in the correct direction.

There were six studies that used a multiple baseline design that were classified as no convincing evidence that the social story was responsible for the behavior change; five of the six studies were unable to show a clear change in the behavior. Five of the six studies had multiple reasons why they were unable to show convincing evidence, including: intervention not being staggered correctly, not showing a clear behavior change, and baseline conditions not trending correctly.

Discussion and Recommendations

The purpose of this paper was to evaluate researchers’ execution of single subject methodologies for studies that implemented social stories for children and adolescents diagnosed with ASD. In total, 41 studies were reviewed. Results of this evaluation indicated that, due to the poor implementation of the various research designs, the vast majority of research studies (92.7%) were unable to offer a convincing demonstration of the effectiveness of social story procedures. Fifty one percent of studies were unable to show any clear demonstration that the social story procedure was responsible for changing participants’ behaviors and 41% of studies were only able to show partial demonstration that the social story procedure was responsible for changing participants’ behaviors. Only 7.3% of studies, through the proper implementation of single subject methodology, were able to show a convincing demonstration that the social story procedure was responsible for changing participant behavior. Thus, the results of this pa-
### TABLE 5

Results for Studies Utilizing Case Study Designs

<table>
<thead>
<tr>
<th>Authors and Year</th>
<th>Number of Participants</th>
<th>Was Social Stories Combined with Other Procedures</th>
<th>Behaviors Targeted</th>
<th>Measurement System</th>
<th>Length of Baseline Condition</th>
<th>Was the Baseline Stable or Trending the Correct Direction</th>
<th>Was the Effect Immediate (within 3 sessions)</th>
<th>Percentage of sessions where the intervention condition overlapped with the baseline condition or behavior was in the wrong direction</th>
<th>Level of Demonstration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bernad-Ripoll</td>
<td>1</td>
<td>Video Modeling</td>
<td>Labeling of emotions, explanation of action responses</td>
<td>Event Recording</td>
<td>10 Sessions</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No: 100%</td>
</tr>
<tr>
<td>Hutchins &amp; Prelock, 2006</td>
<td>2</td>
<td>Comic Book Strips</td>
<td>Being nice towards sibling, stopping insisting or pestering others to continue to play</td>
<td>Subjective Rating Scale</td>
<td>4 Sessions</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No: 18.4%</td>
</tr>
<tr>
<td>Moore 2004</td>
<td>1</td>
<td>No</td>
<td>Sleeping in bed</td>
<td>Estimated Frequency</td>
<td>5 Sessions</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No: 100%</td>
</tr>
<tr>
<td>Noris &amp; Dattilo</td>
<td>1</td>
<td>No</td>
<td>Social interaction (appropriate, inappropriate, None)</td>
<td>Anecdotal</td>
<td>Not Reported</td>
<td>Anecdotal Evidence Only</td>
<td>Not Reported</td>
<td>No: 80.5%</td>
<td>No: No Convincing Evidence</td>
</tr>
<tr>
<td>O’Connor 2009</td>
<td>1</td>
<td>No</td>
<td>Swimming</td>
<td>Anecdotal</td>
<td>Not Reported</td>
<td>Anecdotal Evidence Only</td>
<td>Not Reported</td>
<td>No: No Convincing Evidence</td>
<td></td>
</tr>
<tr>
<td>Okada, Ohtake, &amp; Yangihara 2010</td>
<td>1</td>
<td>No</td>
<td>Elbow Position Morning</td>
<td>10 Second Whole Interval</td>
<td>4 Sessions</td>
<td>No</td>
<td></td>
<td>60%</td>
<td>No: No Convincing Evidence</td>
</tr>
<tr>
<td>Reynhout &amp; Carter, 2007</td>
<td>1</td>
<td>No</td>
<td>Tapping hands</td>
<td>10 S Partial Interval</td>
<td>7 Sessions</td>
<td>Yes</td>
<td>No</td>
<td>58%</td>
<td>No: No Convincing Evidence</td>
</tr>
<tr>
<td>Rogers &amp; Myles, 2001</td>
<td>1</td>
<td>Comic Strip Conversation</td>
<td>Number redirections, minutes tardy</td>
<td>Anecdotal</td>
<td>0 Sessions</td>
<td>Anecdotal Evidence Only</td>
<td>Not Reported</td>
<td>No: No Convincing Evidence</td>
<td></td>
</tr>
<tr>
<td>Swaggart et al., 1996</td>
<td>3 (Only 1)</td>
<td>Using Single Subject Design</td>
<td>Greeting, aggression</td>
<td>Percentage of Opportunity Frequency</td>
<td>9 Sessions</td>
<td>No</td>
<td>Yes</td>
<td>11.1%</td>
<td>No: No Convincing Evidence</td>
</tr>
</tbody>
</table>

- **Number of Participants**: 1, 2, 1, 1, 1, 1, 1, 1, 3 (Only 1).
- **Behaviors Targeted**: Labeling of emotions, explanation of action responses, Being nice towards sibling, stopping insisting or pestering others to continue to play, Sleeping in bed, Social interaction (appropriate, inappropriate, None), Swimming, Elbow Position Morning, Head Position Morning, Head Position Lunch, Head Position Lunch, Tapping hands, Number redirections, minutes tardy, Greeting, aggression.
- **Measurement System**: Event Recording, Subjective Rating Scale, Estimated Frequency, Anecdotal.
- **Length of Baseline Condition**: 10 Sessions, 4 Sessions, 9 Sessions, 4 Sessions, 10 Second Whole Interval, 10 S Partial Interval, 0 Sessions, Percentage of Opportunity Frequency.
- **Was the Baseline Stable or Trending the Correct Direction**: No, Yes, Yes, Yes, No, Yes.
- **Was the Effect Immediate (within 3 sessions)**: No, Yes, Yes, Yes, No.
- **Percentage of sessions where the intervention condition overlapped with the baseline condition or behavior was in the wrong direction**: 100%, 18.4%, 100%, 100%, 60%, 58%, 50%, 11.1%.
- **Level of Demonstration**: No Convincing Evidence.
<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Number of Participants</th>
<th>Design</th>
<th>Was Social Stories Combined with Other Procedures</th>
<th>Behaviors Targeted</th>
<th>Measurement System</th>
<th>Percentage of Baseline Conditions that were stable or trending correctly prior to condition change</th>
<th>Percentage of Intervention Conditions that were stable or trending correctly prior to condition change</th>
<th>Percentage of Intervention conditions where there was a clear change in behavior?</th>
<th>Level of Demonstration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adams et al., 2004</td>
<td>1</td>
<td>ABAB</td>
<td>No</td>
<td>Crying, falling to floor, hitting, screaming</td>
<td>Frequency</td>
<td>37.5%</td>
<td>0%</td>
<td>0%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Agosta et al., 2004</td>
<td>1</td>
<td>ABA</td>
<td>No</td>
<td>Screaming, Sitting</td>
<td>15 S Partial time Interval</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Bledsoe et al., 2003</td>
<td>1</td>
<td>ABAB</td>
<td>No</td>
<td>Spilling food, wiping his mouth</td>
<td>Frequency</td>
<td>0%</td>
<td>100%</td>
<td>50%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Brownell, 2002</td>
<td>4</td>
<td>ABAC</td>
<td>Music</td>
<td>Repetitive statements, instructional repetitions, loud voice</td>
<td>Frequency</td>
<td>62.5%</td>
<td>87.5%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Crozier et al., 2005</td>
<td>1</td>
<td>ABAC</td>
<td>Modified Procedure and Prompts</td>
<td>Talking out</td>
<td>Frequency</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Crozier &amp; Tincani, 2007</td>
<td>3</td>
<td>ABAB, ABAB, ABCAB</td>
<td>Prompts</td>
<td>Sitting, talking to peers, play</td>
<td>Frequency</td>
<td>66.7%</td>
<td>100%</td>
<td>77.7%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Hung, 2011</td>
<td>1</td>
<td>ABA</td>
<td>None</td>
<td>Shouting</td>
<td>Frequency</td>
<td>50%</td>
<td>100%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Ivey, Hefflin, &amp; Alberto, 2004</td>
<td>3</td>
<td>ABA</td>
<td>Prompts</td>
<td>Novel behaviors, Participation Skills</td>
<td>Frequency</td>
<td>67%</td>
<td>66.6%</td>
<td>0%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Knuch &amp; Mirenda, 2005</td>
<td>3</td>
<td>ABA, ABA, ABACA</td>
<td>Reminders</td>
<td>Problem behaviors, cheating</td>
<td>Rate per minute</td>
<td>100%</td>
<td>67%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Kuttler et al., 1998</td>
<td>1</td>
<td>ABAB</td>
<td>No</td>
<td>Precursor to tantrums</td>
<td>Frequency</td>
<td>25%</td>
<td>100%</td>
<td>100%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Lommer et al., 2002</td>
<td>1</td>
<td>ABAB</td>
<td>No</td>
<td>Inappropriate verbalizations, tantrums</td>
<td>Frequency</td>
<td>75%</td>
<td>100%</td>
<td>75%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Mancil, Haydon, Whitty, 2009</td>
<td>3</td>
<td>ABABCBC</td>
<td>No</td>
<td>Screaming</td>
<td>Frequency</td>
<td>80%</td>
<td>53.3%</td>
<td>46.7%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Reichow et al., 2009</td>
<td>1</td>
<td>ABAB</td>
<td>No</td>
<td>Initiations to adults and peers</td>
<td>Frequency</td>
<td>100%</td>
<td>50%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
</tbody>
</table>
## TABLE 7

Results for studies utilizing a Multiple Baseline Design

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Number of Participants</th>
<th>Design</th>
<th>Was Social Stories Combined with Other Procedures</th>
<th>Behaviors Targeted</th>
<th>Measurement System</th>
<th>Percentage of Baseline Conditions that were stable or trending correctly prior to condition change</th>
<th>Percentage of Intervention Conditions that were staggered correctly?</th>
<th>Percentage of Intervention Conditions were there was a clear behavior change?</th>
<th>Level of Demonstration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barry &amp; Burlew, 2004</td>
<td>2</td>
<td>Multiple Baseline across Participants with a Reversal</td>
<td>Prompting</td>
<td>Appropriate Play</td>
<td>Duration</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Chan &amp; O'Reilly, 2008</td>
<td>2</td>
<td>Multiple Probe across Behaviors</td>
<td>No</td>
<td>Raising hand, Social interaction, Vocalizations</td>
<td>Frequency</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>Convincing Evidence</td>
</tr>
<tr>
<td>Delano &amp; Burlew, 2006</td>
<td>3</td>
<td>Multiple Probe Across Participants</td>
<td>No</td>
<td>Appropriate, Inappropriate and No social engagement</td>
<td>Duration and Frequency</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>Convincing Evidence</td>
</tr>
<tr>
<td>Dodd, Hupp, Jewell, &amp; Krohn, 2008</td>
<td>2</td>
<td>Multiple Baseline Design Across Participants</td>
<td>No</td>
<td>Social Skills</td>
<td>Frequency</td>
<td>100%</td>
<td>100%</td>
<td>66%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Gaebe, Mastropieri, &amp; Scruggs, 2009</td>
<td>5 (2 Dropped)</td>
<td>Multiple Baseline Across Participants</td>
<td>No</td>
<td>Time on floor, high pitched voice, hands in mouth</td>
<td>Interval System</td>
<td>66.6%</td>
<td>100%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Hagiwara &amp; Myles, 1999</td>
<td>3</td>
<td>Multiple Baseline Across Settings</td>
<td>Prompting</td>
<td>Washing Hands</td>
<td>Duration, percentage of steps</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Hanley-Hoedorfer, Bray, Kohle, Elinoff, 2010</td>
<td>4</td>
<td>Multiple Baseline Across Participants</td>
<td>No</td>
<td>Verbal Imitations, Contingent Responses</td>
<td>Frequency</td>
<td>75%</td>
<td>0%</td>
<td>0%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Klett &amp; Turan, 2012</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>No</td>
<td>Menstrual Care</td>
<td>Percentage of Steps Correct</td>
<td>100%</td>
<td>50%</td>
<td>33%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Ondemir 2008</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>No</td>
<td>Disruptive Behaviors</td>
<td>15 S Partial Interval</td>
<td>66%</td>
<td>100%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Quibly 2007</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>No</td>
<td>Vocalization, maladaptive behaviors, Bald behaviors</td>
<td>Frequency</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Richter &amp; Test, 2011</td>
<td>3</td>
<td>Multiple Probe Across Participants</td>
<td>No</td>
<td>Correct Responses on a Quiz</td>
<td>Responses</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>Convincing Evidence</td>
</tr>
<tr>
<td>Sansues &amp; Stansfield, 2011</td>
<td>4</td>
<td>Multiple Baseline Across Skills</td>
<td>No</td>
<td>Greetings, Inappropriate humor language, Attempts to touch, Hand in underwear, Videotaping</td>
<td>Frequency</td>
<td>62.5%</td>
<td>0%</td>
<td>25%</td>
<td>No Convincing Evidence</td>
</tr>
<tr>
<td>Sansosti &amp; Powell-Smith, 2006</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>No</td>
<td>Social Engagement</td>
<td>15 S Partial Interval</td>
<td>100%</td>
<td>100%</td>
<td>66%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Sansosti &amp; Powell-Smith, 2008</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>Video Modeling and Prompting</td>
<td>Joining in, Maintaining conversation</td>
<td>15 S Partial Interval</td>
<td>100%</td>
<td>60%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
</tbody>
</table>
per show that, in regards to studies that use single subject methodology, there is little empirical evidence to support that social stories are an effective procedure for children and adolescents diagnosed with ASD. These results are both similar and different from previous reviews on social stories (e.g., Kokina & Kern, 2010). The majority of reviews have indicated that there are methodological limitations to the research evaluating the social story procedure, which limits the conclusions we can make as to the effectiveness of the procedure. Most reviews, however, also go on to state that social stories may be a promising procedure. There appears to be little data to support the claim that social stories are a promising procedure. It is not enough for a research study to show a positive change in behavior; researchers must be able to demonstrate that the positive change in behavior is due to the intervention being evaluated, which is done through proper implementation of a research design. When researchers fail to implement research designs appropriately, it may introduce certain confounding variables, which can limit or qualify interpretation of the results of the study. Unfortunately, this is the case in the first 19 years of social story research. These are two examples of how not adhering to the correct implementation of a research design can mitigate the conclusiveness of a research study to show a positive change in behavior. Researchers must be able to demonstrate that the positive change in behavior is due to the intervention being evaluated, which is done through proper implementation of a research design. When researchers fail to implement research designs appropriately, it may introduce certain confounding variables, which can limit or qualify interpretation of the results of the study. Unfortunately, this is the case in the first 19 years of social story research.

TABLE 7—(Continued)

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Number of Participants</th>
<th>Design</th>
<th>Behaviors Targeted</th>
<th>Measurement System</th>
<th>Percentage of Baseline Conditions that were stable or trending correctly prior to condition change</th>
<th>Percentage of Intervention Conditions that were staggered correctly</th>
<th>Percentage of Intervention Conditions were there a clear behavior change?</th>
<th>Level of Demonstration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scattone, Wilczynski, Edwards, &amp; Rabian, 2002</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>No Disruptive behaviors</td>
<td>10 S Partial Interval</td>
<td>100%</td>
<td>100%</td>
<td>66%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Scattone, Tingstrom &amp; Wilczynski, 2006</td>
<td>3</td>
<td>Multiple Baseline Across Participants</td>
<td>No Appropriate social interaction</td>
<td>10 S Partial Time Interval</td>
<td>100%</td>
<td>50%</td>
<td>66%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Scattone, 2008</td>
<td>1</td>
<td>Multiple Baseline Across Behaviors</td>
<td>Video Modeling Smiling Eye Contact</td>
<td>10 S Partial Time Interval</td>
<td>100%</td>
<td>50%</td>
<td>66%</td>
<td>Partial Evidence</td>
</tr>
<tr>
<td>Schneider &amp; Goldstein, 2010</td>
<td>3</td>
<td>Multiple Baseline Across Behaviors</td>
<td>Visual Schedule</td>
<td>10-15 Sec Momentary Time Sample</td>
<td>0%</td>
<td>50%</td>
<td>66%</td>
<td>No Consimencing Evidence</td>
</tr>
<tr>
<td>Thiemann &amp; Goldstein, 2001</td>
<td>5</td>
<td>Multiple Baseline Across Behaviors</td>
<td>Video Feedback Written Cues Contingent responses, securing attention, comments</td>
<td>Frequency</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>Partial Evidence</td>
</tr>
</tbody>
</table>
of findings. Based on these and other limitations (analyzed in this study), it is not known how effective social stories actually are in changing behavior. Yet, social stories are still commonly implemented (e.g., Reynhout & Carter, 2009) and recommended (The National Autism Center, 2009) for children and adolescents with ASD. The question has to be why is this procedure still being implemented and recommended by clinicians, teachers, and parents. One possible reason why it is being implemented is that it is a relatively easy procedure to implement as compared to more difficult procedures, such as video modeling (e.g., Charlop & Milstein, 1989), script fading (e.g., Sarokoff, Taylor, & Poulson, 2001), and the teaching interaction procedure (e.g., Leaf et al., 2012). A second possible reason is that the social story is still perceived to be an effective intervention (Reynhout & Carter, 2009); this could be due to a lack of understanding of research designs, a self-fulfilling prophecy (e.g., desire for a procedure you are implementing to be successful), or a reliance on subjective measurement.

Based on this review, other recent reviews, and comparative studies that were not included in this analysis (e.g., Leaf et al., 2012), there are several recommendations that can be made for clinicians and researchers. Although social stories is an easy procedure to implement, at this point, clinicians should apply other empirically supported procedures (e.g., video modeling, script fading, behavioral skills training, the teaching interaction procedure) when trying to teach pro-social behavior or decrease aberrant behavior for children and adolescents diagnosed with autism. Although there are a lot of studies that have examined social stories, only three studies have research rigor to show experimental support that the procedure is effective and clinicians should only implement procedures that are demonstrated to be effective in the empirical research.

Although social stories may not currently be a promising avenue for clinicians, teachers, and parents, they are a promising avenue for researchers. Future researchers should evaluate the effects of social stories using a wide variety of single subject methodology; however, these future researchers must ensure that the research methodology is implemented appropriately. In doing so, we can better assess how effective social stories truly are. If empirically sound research shows that social stories are an effective procedure, then future researchers should compare social stories to other procedures (e.g., video modeling, script fading, behavioral skills training) to determine which procedures are the most efficient and effective. In doing so, clinicians will be able to implement the most efficacious empirically proven procedures to individuals diagnosed with autism.

References

* Indicates articles that were evaluated in the review


Received: 5 November 2013
Initial Acceptance: 6 January 2014
Final Acceptance: 20 February 2014
Compensation Age Theory: Effect of Chronological Age on Individuals with Intellectual Disability

Hefziba Lifshitz-Vahav
Bar-Ilan University

Abstract: The main goal of this article is to discuss a new concept, the “Compensation Age Theory (CAT)”, for individuals with intellectual disability (ID). The CAT is a holistic framework comprised of four dimensions: (a) the state of the art of the CAT; (b) the theoretical resources which are at the core of the CAT; (c) a series of empirical studies performed by the author and other scholars which support the assumptions of the CAT; (d) cognitive educational intervention programs for individuals at all levels of ID throughout the lifespan, which are anchored in this theory. The CAT framework supports the view that all human beings, even people with ID, are capable of change. Despite the limitations imposed by age or severity of disability, the concept "self-actualization" can be expanded to include people with ID who, under appropriate environmental conditions and with continuous systematic intervention, can be brought to a level of functioning previously absent from their behavioral repertoire.

The main goal of the present article is to discuss a new concept, the “Compensation Age Theory (CAT)” for individuals with intellectual disability (ID). The CAT is a holistic framework comprised of four dimensions: (a) the state of the art of the CAT; (b) the theoretical resources which are at the core of the CAT; (c) a series of empirical studies performed by the author and other scholars which support the assumptions of the CAT; (d) cognitive educational intervention programs for individuals at all levels of ID throughout the lifespan, which are anchored in this theory.

CAT-State of the Art

When considering cognitive education or cognitive intervention programs for individuals with ID, it is assumed that the weight of the mental age (MA) or the basic cognitive level is the crucial factor for determining their cognitive ability. The CAT postulates that the weight of the chronological age (CA) is similar to that of the MA, and that CA plays an important role in determining the cognitive ability of individuals with ID beyond their MA. The CAT claims that in later years, there is compensation for the developmental delays experienced by individuals with ID in their early years. Furthermore, the cognitive ability of adults with ID can be modified at a critical period which was considered to be non-modifiable, even at an advanced age. Our argument is that years of schooling, greater maturity and cumulative life experience help them acquire skills that were previously absent from their behavioral repertoire.

The CAT emphasizes the effect of CA on the cognitive ability of individuals with ID. Cognition is a wide concept which involves several terms including cognitive skills, cognitive thinking, cognitive strategies, self-regulation skills and metacognitive processes (Hessles & Hessels-Schlatter, 2013; Feuerstein & Falik, 2010; Haywood, 2013; Vygotsky, 1978). Ashman and Conway (1997) defined cognition as a sequence of mental capacities such as "storing, retrieving, storing, transforming and manipulation of information" (p. 41). Feuerstein and Falik (2010) added other skills, such as “organization, processing, understanding generalization, transferring and transformation”.

There are several reasons for emphasizing
cognition by the CAT in individuals with ID. (a) In line with the Structural Cognitive Modifiability Theory (Feuerstein & Rand, 1974; Feuerstein, 2003), the CAT claims that individuals with ID can benefit not only from exposure to concrete information and from sensory-based experiences, but also from programs designed to ameliorate impaired cognitive functions and to enhance abstract thinking (Lifshitz & Rand, 1999). (b) In line with the Normalization (Wolfensberger, 2002) and Quality of Life (Schalock, 1996) principles, the CAT is guided by the vision that all individuals with disabilities will be valued members of their community and society. Individuals with ID at all levels should have meaningful participation in all aspects of life as stated in the UN Convention on the Rights of Persons with Disabilities (2006): “Persons with disabilities continue to face barriers in their participation as equal members of society and violations of their human rights in all parts of the world” (p. 4). Our purpose as a society is to “promote their participation in the civil, political, economic, social and cultural spheres with equal opportunities” (p. 4). Participation in the above activities requires manipulation of the above-mentioned cognitive skills and strategies. (c) We live in a digital age and are surrounded by computers and other technological instruments such as smartphones and iPads which influence fundamental areas of everyday life (Shamir, 2013). These instruments are also accessible to individuals with ID. Coping with the rapidly changing technological world, demands activation of the above cognitive skills. (d) Accelerated progress in brain and neuro-psychological science indicates a reciprocal relation between brain and cognitive modifiability. As will be shown below, fMRI experiments indicate neural plasticity of the brain as a result of a cognitive-stimulating environment even among individuals with ID (Head, Lott, Patterson, Doran & Haier, 2007).

CAT-Theoretical Background

One of the intriguing questions in the professional literature that deals with individuals with ID is: how do intelligence and cognitive functioning develop and decline throughout their lifespan? Three possible models of intelligence or cognitive trajectories in individuals with ID compared to the general population have been proposed over the years (Fisher & Zeaman, 1970, and others): the Impaired Trajectory, the Parallel Trajectory and the Compensation Trajectory. These models are based on traditional theories of intelligence in the general population (Wechsler, 1955; Kauffman, 2001), according to which there is an increase in intellectual functioning up to the age of 20, and that the increase in intelligence (the general intelligence score without a distinction between crystallized and fluid measures) is linear (development). This is followed by an asymptote (stability), and a decline from around the age of 60. The three models of cognitive trajectories among individuals with ID differ in three dimensions: (a) the age at which intelligence reaches its peak; (b) the length of the stability or asymptote period; (c) the age of the onset of decline. The three possible trajectories of intelligence and cognitive growth and decline in individuals with ID are presented in Figure 1. We will present the three models, with emphasis on the third model which comprises the basis for the CAT theory.

The “Impaired Trajectory” (IT) model is anchored in the Cognitive Reserve (CR) Theory (Katzman, 1993), which posits that normally occurring individual differences in the way people process tasks might provide a differential reserve against brain pathology or age-related changes. A sub-concept of the CR theory is the “neural reserve”, which represents the capacity to perform tasks or cope with increasing task difficulty. Individual differ-
ences may result from innate characteristics (e.g., intelligence), or may be modulated by life events such as educational or occupational experiences or leisure activities (Scarmeas & Stern, 2003; Wilson & Bennett, 2005). A higher neural reserve might be expressed in the form of brain networks that are either more efficient or have a greater capacity to handle increased demand. Individuals with ID possess less CR than their peers without ID, by definition (Zigman et al., 2004). They have lower intelligence and slower task processing. They do not achieve higher educational levels or occupational statuses, and tend to participate in fewer intellectually stimulating leisure activities. The IT model predicts that individuals with ID will exhibit restriction in developing intelligence before their 20’s. Thereafter, they will exhibit stability. Their lower CR exposes them to an accelerated and steeper decline which will be manifested in their mid-thirties or forties.

The Parallel Trajectory (PT) model was proposed following recent studies that questioned the applicability of the CR theory to individuals with ID (Oliver, Crayton, Holland, Hall, & Bradbury 1998). If people with ID have reduced CR compared with their peers with TD, by definition, then these persons would be expected to be at greater risk for dementia of the Alzheimer type (DAT) with increasing age than the general population (Snowdon, Greiner, & Markesbery, 2000). However, several studies (Merrick, 2010; Zigman et al., 2004) found an equivalent or even lower risk of dementia among adults with ID with a nonspecific etiology (NSID). Based on the lower age for onset of DAT among individuals with ID, the above authors reached the conclusion that factors that determine intelligence may have little or no direct relationship to risk for dementia in individuals with NSID. On the contrary, studies indicate similar trends of cognitive change among adults with ID and adults with typical development (TD).

Facon (2008) supported the similar trajectory model using WAIS-R results of adults with and without ID aged 20 to 54 years. Hierarchical regression analyses among the two groups showed similar evolution of scores with increasing age for verbal and performance scales. Devenny, Hill, Patxot, Silverman, and Wisniewski (1992) reported similar trends of change compared to the general population among adults with ID with and without Down syndrome (DS) aged 50+ in short and long-term verbal memory and visual-spatial organization. Based on the above, the PT model predicts that intelligence and cognitive development among individuals with ID are similar to those of the general population in terms of the peak age and the onset and rate of decline. The difference between the two groups lies in the baseline IQ level, which in individuals with ID is two standard deviations below the norm of the general population.

The Compensation Trajectory (CT) model, which is at the core of the current study, postulates that while intelligence in the general population grows linearly up to the age of 20, individuals with ID will be compensated in later years due to the developmental delays they experience in their early years (Fisher & Zeaman, 1970).

The proposed CAT theory distinguishes between the period of cognitive growth and the period of cognitive stability and decline. It predicts that the duration of the growth period among individuals with ID may be longer compared to the general population, and may continue until their late 40’s. Intelligence and cognitive functioning in this population will therefore reach their peak at the age of forty to fifty. The theory also predicts that similarly to the general population, intelligences among individuals with ID will exhibit stability between the ages of 50 to 60, and will show a decline after age 60. Thus, individuals with ID will exhibit a different pattern of cognitive growth than individuals with TD, but will show a parallel pattern of cognitive stability and decline.

The CAT is anchored in several theoretical resources: the Structural Cognitive Modifiability (SCM) theory and the Active Modifying (AM) approach (Feuerstein & Rand, 1974; Feuerstein, 2003; Feuerstein & Falik, 2010). The basic assumption underlying this theory is that the human organism is, by nature, a system open to its environment and accessible to change, even in the presence of three formidable obstacles usually believed to prevent change: (a) age; (b) etiology; (c) severity of limitation (Feuerstein, 2003).

Theoreticians such as Bloom (1964) and Piaget (1970) argued that the age of an indi-
individual is a crucial factor for the success of environmental intervention. In their opinion, the critical time to carry out an intervention aimed at correcting cognitive deficiencies is at a young age, and the sooner the intervention the better the results. The SCM theory does not dispute the importance of starting intervention programs at an early age. However, it rejects the assumption that there is a critical time for initiating such an intervention. This position is buttressed by Gunzburg (1968) and Luftig (1987), who argued that maturation and life experience of adults with ID help them learn and acquire new cognitive skills that were previously absent from their behavioral repertoire.

The Cognitive Reserve Theory (CRT) (Stern et al., 2005) is also at the foundation of the CAT, as is its sub-concept, neural compensation. This concept relates to a situation in which the physiological effects of aging or brain pathology cause the alteration of a brain network, resulting in a network that would not normally be used by unaffected individuals but in a different way, or additional brain areas might be recruited. Stern et al. hypothesized that the altered network is used to compensate for the inability to utilize the healthy brain’s responses to increased task difficulty. The ability of the compromised brain to express or optimize compensatory networks may also vary as a function of CR.

One might argue that individuals with ID are exposed to lower CR due to their lower level of intelligence, fewer opportunities for cognitive education and cognitive leisure activities. Based on the lower rate of dementia among the population with ID compared to the general population (Zigman et al., 2004), our argument is that CR in individuals with ID should be examined within the population with ID itself, and not compared to the general population. There are individual differences within the population with ID in task processing, according to the intelligence level and life events such as occupational and leisure experience. However, as will be shown, individuals with ID can acquire cognitive skills efficiently in their adulthood and even at an advanced age. This statement includes all levels of ID: mild/moderate as well as severe/profound.

Developing cognitive skills during young adulthood, in midlife, is part of the agenda of the CAT. It is anchored in the Cognitive Activity Theory (Wilson & Bennet, 2005). This theory postulates that there is an association between participation in cognitive activities during midlife, and a reduced risk of cognitive decline leading to Alzheimer disease (Stern et al., 2005; Wilson & Bennett, 2005; Wilson, Barnes, & Bennett, 2007). Based on the Cognitive Reserve Theory (Stern et al., 2005), the authors suggest that cognitively active persons experience more cognitive decline before reaching the cognitive impairment levels commensurate with dementia, compared with less cognitively active adults.

This association between cognitive activity and reduced risk of Alzheimer disease has been confirmed in several studies (Scarmeas, Levy, Tang, Manly, & Stern, 2001; Winocur, Palmer, Dawson, Binns, Bridges, & Stuss, 2007). In one study (Wilson & Bennett, 2005), healthy older Catholic clergy members rated how frequently they participated in several cognitively stimulating activities at the beginning of the study. After five years, persons reporting frequent participation in cognitively stimulating activities had only half the risk of developing Alzheimer disease compared with those reporting infrequent cognitive activity. The authors suggest that education affects the risk of cognitive impairment and dementia by somehow enhancing the brain’s capacity to tolerate Alzheimer disease pathology.

One might argue that cognitive activities include a high level of abstract thinking which are beyond the capacity of individuals with ID. However, according to Wilson and Bennett (2005), even activities that involve a basic level of information processing can be fruitful. In fact, all kinds of leisure activities, such as listening to the radio, watching TV, reading, writing, playing games, fishing, leisure creations, involve cognitive stimulation. Thus, activating cognitive interventions among adults with ID at all levels of ID might be a protective factor against accelerated deterioration.

**CAT-Empirical Findings**

Our claim about a compensation age and our concentration on adulthood and even advance age is based on empirical findings. The first assertion about the influence of CA on the cognitive ability of individuals with ID was
presented by Facon his colleagues. Facon and Facon-Bollengier (1997) examined the effect of CA on the crystallized intelligence level of children and adolescents (aged 6–18) with ID. Participants were tested on both the Peabody Pictures Vocabulary test (Dunne & Dunne, 1981) and the French version of the Columbia Mental Maturity Scale. Stepwise regression showed that 28% of the PPVT in the above sample was accounted for by CA. In another study (with the same ages), Facon and Facon-Bollengier (1999) showed that the fluid intelligence factor and CA explain an important fraction of the crystallized intelligence factor variance (43% and 21%, respectively) of individuals with ID. According to Facon and his colleagues, this finding supports the hypothesis that CA-related experiences exert a significant effect on the crystallized component of intelligence in people with ID. The authors claim that long life experience may aid people with ID to succeed better in some cognitive tasks and partly determines their MA. Thus, when subjects are matched by MA, as is usually done to separate the influence of CA and MA on target tasks, there is a de facto attenuation of the contribution of the CA.

The role of CA on the cognitive ability of individuals with ID is also expressed by the criteria of matching between individuals with ID and with TD in memory studies. When individuals with ID and TD were matched according to CA, individuals with TD outperformed those with ID, a finding which is self-explanatory (Turner, Hale, & Borkowski, 1996; Cherry, Njardvik, & Dawson, 2000). However, when matching was based on MA, the performance of individuals with ID was equal to their controls. In this case the individuals with ID are older than their TD peers and this affords them an advantage compared to their controls (Carlin, Soraci, Dennis, Chechile, & Loiselle, 2001; Perrig & Perrig, 1995, and others). Carlesimo, Marotta, and Vicari (1997) stated that longer exposure to linguistic and academic experiences of adolescents with ID may explain their more efficient use of semantic strategies than their TD peers with the same MA. While those studies focus on younger ages (children/adolescents), a series of studies conducted by the author of the current study implies the influence of CA on the cognitive ability of adolescents and adults with ID. These studies are presented henceforth in brief:

Study 1 (Lifshitz & Tzuriel, 2004) was performed with the goal of examining the influence of a cognitive intervention program on cognitive achievements of adults with ID and the duration of this influence. The original sample (Lifshitz & Rand, 1999) included 71 adults with ID in three age groups: young adults with ID (CA 20–35); middle-age (36–50); older adults (CA 50–70) with IQ 40–70. The follow-up study conducted three years later included 21 adults (CA 30–59) who participated in the original study.

The central means of intervention were four tools from the Instrumental Enrichment Program (Feuerstein, Rand, Hoffman, & Miller, 1980): comparison, categorization, time and space relations. The effects of the intervention were examined with reference to three types of thinking (Glanz, 1989): logical thinking (Reversal Test and Test of Verbal Abstraction), predictive thinking (Maze Tests) and insightful thinking (Postures Test and Children Test). The battery was administered five times: twice before the intervention spaced two months apart, twice afterwards spaced two months apart, and a follow-up three years later.

MANOVAs and contrast between the four time points in the original study yielded significant improvement from Time 2 to 3, and two months later (Time 4), a divergency effect in logical and predictive thinking. The MANOVAs in the follow-up evaluation (Time 5) showed a drop in the cognitive functioning relative to Time 4, but participants maintained their achievement in relation to Time 3, a finding which indicates a durability effect. Figure 2 presents the cognitive achievements of 21 participants in three types of thinking in the five time points.

The findings show significant changes among all investigated age groups (30-57), including adults with ID with and without Down syndrome, without any differences due to age. The results indicate that it is possible to alter the level of cognitive functioning of adults with ID even at an advanced age.

Study 2 (Lifshitz, Tzuriel, Weiss, & Tzemach, 2010) was performed in order to map the difficulties and cognitive processes in solving analogical problems among adolescents (age...
13–21) and adults (age 25–66) with ID (IQ 40–70). The Conceptual and Perceptual Analogical Modifiability test (Tzuriel & Galinka, 2000) was administered using a dynamic assessment procedure.

Figure 3 presents the improvements from pre to post-teaching tests between adolescents and adults. Repeated measures MANOVAs and post-hoc tests did not reveal significant differences between the two age groups in the pre-teaching stage. In the post-teaching stage, the adult group scored significantly higher ($p < .05$) than the adolescent group ($M_{adolescent} = 15.09$ and $M_{adult} = 8.28$, respectively). The findings indicate that the adults gained more from teaching in the DA procedure than the adolescents.

Study 3 (Lifshitz & Katz, 2009) was performed in order to examine: (a) the level of understanding of Jewish cognitive concepts among Jewish adolescents and adults with ID (IQ 40–70); (b) the psychological emotional motives of the participants for being religious. The participants included adolescents (age 13–21) and adults (age 30–60) with ID. The cognitive component included four factors: Concept of God, Heavenly recompense-reward and punishment, Divine providence - the sense of the presence of God in everyday life, efficacy of prayer.

The scores of prayer efficacy and providence of God were significantly higher among the adults than the adolescents (Figure 4). The adults also exhibited more mature motives for being religious (dependence, security, God as an anchor) than the adolescents (belonging to a religious community, expect to fulfill personal wishes). Regression analysis indicated that among the adolescents, MA contributed to the explained variance of the cognitive components, while among the adults, CA contributed to the explained variance of the cognitive component.

Study 4 (Lifshitz, Klein, & Fridel, 2010) was carried out in order to examine the effects of a year-long Mediational Intervention for Sensitizing Caregivers (MISC; Klein, 1992) on the quality of interactions between rehabilitation day center paraprofessionals and their adult consumers with severe ID. Another goal was to examine the effect of the intervention on the consumers’ cognition, autonomy, and behavioral functioning.

The objective of the MISC (Lifshitz & Kein, 2007) is to help caregivers and direct staff...
relate to their dependents in a way that will enhance their cognitive, autonomous, and behavioral functioning. It is not content-specific, but may serve as a tool for teaching “literacy of interaction” in daily activities such as vocational, domestic, and leisure skills.

Paraprofessional staff members in rehabilitation centers and their consumers with severe and profound ID were divided into an experimental and a control group. The paraprofessionals in the experimental group participated in a workshop on the MISC and then activated the MISC intervention in the rehabilitation center for one year.

Following the intervention, more mediation of choice making, cognitive expansion, and encouraging with explanation were observed among the paraprofessionals in the experimental group than in the control group. Consumers with ID in the MISC group improved their arithmetic skills, temporal concepts, and sequential memory of two digits (Figure 5). The findings indicate that appropriate environmental conditions and continuous systematic intervention may enable adults with severe and profound ID to invent new skills, which were previously absent from their behavioral repertoire.

The idea of a compensation mechanism with increasing age in individuals with ID is also supported by genetic and brain studies conducted among adults with DS. It is known that adults with DS develop Alzheimer disease pathology progressively with age (Janicki & Dalton, 2000; Oliver et al., 1998), but clinical signs of dementia are delayed by at least 10 years after the first signs of disease. Furthermore, studies have reported a subset of adults with DS who do not exhibit dementia at any age (Devenny et al., 1992; Zigman et al., 2004). While virtually all DS subjects older than 40 have a significant neuropathology of dementia, there is a lack of concordance between the typical age of onset for dementia in this etiology (Mesulam, 1999). Head et al. (2007) suggested that compensatory events may be of particular relevance for the DS group. Their claim is based on fMRI and PET observations on individuals with DS which indicated compensatory increases in the metabolic rate in vulnerable brain regions in DS prior to the development of dementia. Head et al. (2007) suggested that genes which are overexpressed in DS (APP, DSCAM, MNB/DYRK1A, RCAN1) produce proteins critical for neuron and synapse growth, development and maintenance and provide further evidence for the activation of plasticity mechanisms in this etiology. These genes may lead to developmental cognitive deficits, but paradoxically, with aging may participate in molecular cascades supporting neuronal compensation. Based on work in rodent models and in a canine model of human brain aging, Head et al. suggested that use of behavioral enrichment (including physical exercise) may have a significant impact on healthy brain aging in DS. These same interventions may promote pathways and molecular cascades involving genes overexpressed in DS that may enhance compensatory mechanisms.

**CAT: Educational Implications –four programs of cognitive education for individuals with ID**

The fourth dimension of the CAT framework is its operative implications for cognitive educational programs. We constructed four cognitive educational intervention programs for individuals at all levels of ID: (a) Empowerment program: Academic enrichment for adults with mild/moderate ID; (b) ABC: Enriching cognition and literacy, affect and behavior skills during daily life activities for school-age students with severe/profound ID; (c) ABC: Enriching cognition and literacy affect and behavior skills during daily life activ-
ities for adults with severe/profound ID; (d) ABC: Enriching cognitive and literacy affect and behavior skills during daily life activities for elderly persons with ID and Alzheimer disease.

(a) Empowerment program: Academic enrichment for adults with mild/moderate ID: This program is designed for adults with moderate and mild ID with/without DS. The program combines humanistic and scientific goals. It emerges from the UN convention for persons with disabilities: "Parties shall ensure an inclusive education system at all levels and lifelong learning directed to: The full development of human potential, sense of dignity and self-worth, their talent creativity as well as their mental and physical abilities" (UN, 2006, p. 20). The empowerment program puts these rights into action.

In this program, the students attend the School of Education, Bar-Ilan University (during the academic year) once a week for six academic hours. The courses taught are psychology, sociology, self-advocacy, library and computers and are adapted to the level of the students. The lecturers are students in the Intellectual Disabilities track of the Master’s Degree program at the School of Education. Teaching in the project is part of the practicum in the track.

Another group of students with ID are included during the academic year in a BA research seminar on Lifelong Learning of Individuals with Disability, together with regular students. One goal of this special seminar is to teach students with ID to conduct research about themselves. Reciprocal learning takes place between regular students and students with ID. Together they study issues related to self-advocacy. The task of the students with ID is to interview three friends with ID about self-concept, self-efficacy, hope and optimism using questionnaires. The typical students perform the statistical analysis. The two groups analyze the results and draw conclusions together.

The educational objectives of these programs are to acquire knowledge on academic subjects that might be relevant to this population, develop strategies for learning, access the university’s libraries, conduct small research projects and use the computer lab. The social objectives are to expose students with ID to students with TD in class and during breaks, expand the friendship circle of students with ID, empower and strengthen their self-image, confidence, and quality of life, and construct positive attitudes towards individuals with disability among the regular students.

(b) ABC: Enriching cognition and literacy affect and behavior skills during daily life activities for school-age students with severe/profound ID (Lifshitz-Vahav, Tal, Nissim, & Nissim, in press): This program was developed by us as a request from the Special Education Division of the Israeli Ministry of Education and serves as the new Israeli national curriculum for students with severe/profound ID. The essence of this program (published earlier as the MISC approach, Klein, 1992; Lifshitz & Klein, 2007; Lifshitz, Klein, & Fridel, 2010) is introducing cognition, literacy and autonomy during daily life activities to school-age students with severe and profound ID. The program includes three components: behavioral component (mediation of adaptive behavior skills), cognition (mediation of cognition and literacy), affect (mediation of autonomy). Mediation of these components is conveyed by the mediators (teachers, paraprofessionals and direct caregivers) through their interaction with their students with ID not only in the formal lessons but during meal time, domestic skills, sport class, art, occupational therapy, vocational preparation, etc.

There is a myth among teachers, paraprofessionals and even parents that literacy and cognitive skills are beyond the ability of individuals with severe and profound ID. Our experience with the ABC (the MISC) intervention among adults with severe and profound ID (Lifshitz et al., 2010) indicates that individuals with severe and profound ID can be modified even at advanced age. Based on the above, we constructed the program for school-age students (age 6–21) with severe and profound ID. Some examples for introducing cognition during daily life activities are presented: alongside the impartation of how to use the spoon or fork, the mediator can talk with the students about the offered foods, the taste, the color (this tomato is red), the shape (round). When eating a cake, the mediator talks about the ingredients, the nutritional components. Word signs of the menu are indicated. In art class, teachers talk about the
color, shape, design, texture, verbs and nouns related to the work.

Autonomy is integral part of human rights and affords meaning to life (affect), as stated in the UN convention for persons with disabilities (2006): “Recognizing the importance for persons with disabilities of their individual autonomy and independence, including the freedom to make their own choices” (p. 3).

For individuals with severe/profound ID, the principle of making choices should be applied in everyday activities such as beloved foods, choosing their clothes, the occupational and leisure activities. The three components of the ABC, including the strategies and examples, are presented in Table 1.

(c) ABC: Enriching cognition and literacy, affect and behavior skills during daily life activities for adults with severe/profound ID (Lifshitz et al., 2010): This program was designed for adults with severe/profound ID. There have been attempts to improve cognitive, choicemaking, and adaptive behavior skills of adults with severe and profound ID. However, these studies taught participants in separate classes or on an individual basis (1:1 ratio between mediator and trainee). These studies focused on specific skills: improving metacognition by a computer-assisted program (Moreno & Saldana, 2005), receptive communication abilities (Casella, 2004), choosing the leisure activity (Browder, Cooper, & Lim, 1998), etc., and lacked a holistic approach to combine all these skills together in everyday natural settings. Our program advocates the ABC, a holistic and broader approach to concurrently improving the cognition and literacy, behavior and affect of individuals with severe ID through daily life activities via their ongoing and varied interactions with their paraprofessional staff. This program was adapted for life situations of adults and was implemented for one year in two vocational centers. The findings indicate, as mentioned above, that before the interaction, the direct staff in the control group continued to emphasize only the use of basic skills for immediate performance, without attempting to expand on their meaning by giving basic information such as color, size, or number of products. For example, during vocational work, consumers were instructed to pack 10 spoons in plastic bags and were provided with a box containing 10 slots to aid them in counting the spoons. Staff members guided them in inserting the spoons in the slots. In the experimental group, the paraprofessionals counted aloud with consumers and then provided them with opportunities to count alone, but staff in the control group did not. Following the intervention, more mediation of autonomy (choice making) was observed in the experimental group. Through their interactions they succeeded in improving math skills and time orientation.

(d) ABC: Enriching literacy, affect and behavior skills during daily life activities of persons with ID and AD (Lifshitz & Klein, 2011): One of the serious problems resulting from the increase in lifespan is a concomitant rise in cases of dementia of the Alzheimer type (DAT) among adults with/without DS. Estimates of the age-specific prevalence of dementia in adults with DS have varied widely, from under 10% to over 75% (Zigman, Schupf, Sersen, & Silverman, 1996). This association appears to be due to a triplication of the gene for the beta-amyloid precursor protein (β-APP) which is located on the proximal part of the long arm of chromosome 21 (Goldgaber, Lerman, McBride, Saffiotti, & Gajdusek, 1987).

Several studies focused on the stressors and sense of burden of the caregivers (families) or staff caring for elderly persons with ID and AD (McCallion, McCarron, & Force, 2005). Caregivers lack knowledge on how to treat and what can be done with persons with ID who are exposed to Alzheimer disease, and exhibit problems with memory, space and time orientation. The ABC for adults with ID and AD affords a solution to these questions and fills this void. The mediational parameters can be applied through the interaction between caregivers and persons with ID/AD during daily activity: meal and medication time, work sessions and leisure activities. This ABC for adults with ID and AD is based on the integration of person-centered cultural approaches which emphasize individual identity and selfhood and the cognitive rehabilitation approach (Clare, Wilson, Carter, & Hodges, 2003) which is based on the understanding that despite difficulties with memory and other cognitive functions, people with dementia still have the ability to learn new associations and information, and to adjust their behavior and responses. In line with the CAT and the cognitive rehabilitation ap-
### TABLE 1

**Three mediational components of the ABC**

<table>
<thead>
<tr>
<th>Cognitive component</th>
<th>Cognitive expansion:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Giving meaning to all activities such as work, leisure activities, ADL. Creating a need to search for meaning in the individual’s experience to maintain the attention gained in the focusing process and allowing for the mediating process.</td>
</tr>
<tr>
<td></td>
<td>Clarifying processes (insight); attributing past and future needs to the present situation; critical interpretation; providing general rules, inductive/deductive reasoning; spatial and temporal orientation; providing appropriate strategies for memory and recall.</td>
</tr>
<tr>
<td></td>
<td>Relating to concepts such as colors, size, shape, numbers, counting and learning a sequence of actions of the vocational tasks, meals, and leisure activities.</td>
</tr>
<tr>
<td></td>
<td>Giving additional information about work beyond the immediate experience. Using word signs accompanied by pictures in the vocational hall, dining room (menu, food’s ingredients), and leisure activities (work tools, actions, materials).</td>
</tr>
<tr>
<td></td>
<td>Focus in ADL – Feeding: the names of foods, menu, food ingredients should be indicated by word signs accompanied by pictures.</td>
</tr>
<tr>
<td></td>
<td>Vocational activity: Focusing on the object to be packaged, weighed, sorted. 10 items relevant to work such as the name of products should be indicated by word signs.</td>
</tr>
<tr>
<td></td>
<td>Space orientation: The close surroundings: room, shower, toilet, dining room, infirmary, should be indicated by word signs accompanied by pictures.</td>
</tr>
<tr>
<td></td>
<td>Providing choices for decision-making appropriate for all levels of ID including individuals with severe/profound ID during performing ADL skills, vocational, leisure activity. (choosing the type of work, the loved food, the leisure activity)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Affect component</th>
<th>Mediation of meaning (affecting):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enhancing quality of life by providing opportunities for choice making, self-determination, and autonomy (and other values).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavioral component</th>
<th>ADL skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Includes all stages of task analysis (Gold, 1978; Luftig, 1987) when performing each task.</td>
<td></td>
</tr>
<tr>
<td>B. Behavior that models, demonstrates and/or verbally suggests regulation of behavior in relation to specific tasks or any other cognitive process required before overt action is introduced.</td>
<td></td>
</tr>
</tbody>
</table>
approach, the ABC for adults with ID and AD works on mediation of cognition and autonomy during daily life activities in order to improve their functioning in the same areas that they exhibit deficit (for more details see Lifshitz & Klein, 2011).

For example, Jacob (54 years old) is a person with DS who was diagnosed with early-stage Alzheimer disease four years ago. Based on the cognitive rehabilitation theory (Clare et al., 2003), a tailored program with three mediational parameters of the MISC was constructed in order to overcome his weaknesses. The findings indicate that he showed numerous important strengths, including the capacity and motivation for learning new skills. This suggested that he was able to learn new strategies that compensate for his deterioration in short-term memory, orientation in time and space. Jacob was coping with the onset of dementia by facing up to its impact and trying to adapt.

Conclusions

The CAT framework supports the view that all human beings, even people with ID, are capable of change. The CAT model is not just a theoretical framework. It can be implemented in the field of cognitive education among individuals with ID, especially in older ages and among those with severe and profound ID. Individuals with ID can benefit not only from exposure to concrete information and sensory-based experiences, but also from programs designed to ameliorate impaired cognitive functioning. Despite the limitations imposed by age or disabilities, the concept "self-actualization" can be expanded to include individuals with ID at all levels, even at an advanced age.

Limitations and Further Research

The CAT focuses on the effect of CA on the cognitive ability of individuals with ID. Except for one research (Lifshitz & Katz, 2009), it does not relate to emotional and social aspects. It is recommended to expand research on emotional and social aspects as well. The empirical findings of the CAT were based on behavioral studies. Conducting brain and neuro-imaging studies will strengthen the CAT from the neurological point of view.

References


Received: 25 November 2013
Initial Acceptance: 28 January 2014
Final Acceptance: 28 May 2014
Incorporating Functional Digital Literacy Skills as Part of the Curriculum for High School Students with Intellectual Disability

David F. Cihak and Rachel Wright
University of Tennessee

Cate C. Smith
Appalachian State University

Don McMahon
Washington State University

Kelly Kraiss
University of Tennessee

Abstract: The purpose of this study was to examine the effects of teaching functional digital literacy skills to three high school students with intellectual disability. Functional digital literacy skills included sending and receiving email messages, organizing social bookmarking to save, share, and access career websites, and accessing cloud storage to download, revise, and upload documents. Results indicated that all students acquired and maintained these functional digital literacy skills. Findings are discussed in the context of teaching essential digital literacy skills to increase greater participation in a digital society.

As the technology wave continues to permeate every aspect of daily life, many organizations struggle to evolve accordingly, and education is no exception. The rapid rise in the use of technology has significantly impacted the way society functions, and educators face many new challenges in teaching students of the digital age. These challenges include the need to reevaluate and redefine previously agreed upon concepts, such as what it means to be literate in today’s digital world. The term “literacy” must be expanded beyond its traditional meaning of the ability to read and write to reflect the skills needed to interpret and understand information presented through the existing realm of digital media. The basic principles of traditional literacy instruction must also adapt to match advancements in concepts of literacy accordingly. Adoption of such digital literacies skills is essential for educators to not only keep their students engaged and motivated, but to better prepare them for a technologically-oriented workplace (Collier, 2007). Students with intellectual disability (ID) face this same future, and the same demands of the technologically-oriented workplace, and therefore must be provided explicit instruction that will promote independence and increase opportunities in adulthood.

For students with ID, traditional literacy instruction has primarily focused on teaching sight word recognition and is often isolated from meaningful context (Clendon & Erickson, 2008). Despite what is known about the components of high quality reading instruction, reading instruction for students with ID has traditionally emphasized functional reading skills such as sight words necessary for daily living, safety, and independence as a singular, principal focus (Browder, Wakeman, Spooner, Ahlgrim-Delzell, & Algozzine, 2006; Kiewer, 1998). Shifting perspectives of special educators and improvements in legislature (IDEA, 2004; NCLB, 2002) indicate a more expansive and potentially liberating view of literacy and learning for students with ID. As a result of these changes, educators and proponents of students with disabilities have advocated for access to an instructional program that would promote participation and progress in the general curriculum, including literacy instruction (Jackson, 2005). Although...
sight words are an essential component of literacy instruction, comprehension and communication are key skills for students with ID (Browder et al., 2006). A body of research demonstrates that students with ID can successfully acquire, maintain and generalize new literacy skill with systematic prompting and feedback (e.g., Browder, Ahlgrim-Delzell, Spooner, & Baker, 2009; Browder, Lee, & Mims, 2011). Multiple literature reviews establish the use of systematic instruction as an evidence-based practice for teaching academic content to students with ID (Browder et al., 2006; 2009; 2011; Spooner, Knight, Browder, & Smith, 2012).

Alberto et al. suggested that literacy skills provide individuals with the “ability to obtain information from the environment with which to make decisions, alter the environment, and gain pleasure” (Alberto, Frederick, Hughes, McIntosh, & Gihak, 2007, p. 234). Students with ID must learn strategies that enable them to access the information available from the environment, which may be presented in a variety of media formats in addition to printed text. By incorporating elements of visual literacy (the ability to discern meaning conveyed through images) into the functional literacy curriculum for students with moderate intellectual disabilities, Alberto et al. demonstrated that explicit, systematic instruction of picture and logo reading is an effective strategy to increase students’ abilities to interpret meaningful information from the environment. Explicit teaching of basic visual literacy skills provides instruction that is not only functional, but also allows access to information available in multiple settings, including the real-world and online communities. As digital information continues to dominate our lives, visuals and graphics created with new technologies are becoming as common as digital text, and visual literacy is critical for access to and participation in digital environments and communications.

Technology provides students with ID with numerous opportunities to access information that they may not have been able to obtain otherwise. Many types of technological devices are available to teach a variety of types of skills to students with ID, including academic, social-communicative, community, and vocational skills (Browder et al., 2006; Gihak, Kesler, Alberto, 2007, 2008; Davies, Stock, & Wehmeyer, 2002a, 2002b, 2004; Ferguson, Myles-Smith, & Hagiwara, 2005). Devices that have been used successfully include computers (e.g., Coleman, Hurley, & Gihak, 2012; Ramdoss et al., 2011), handheld computers (e.g., Gihak, Wright, & Ayres, 2010; Mechling, Gast, & Seid, 2010), iPads (e.g., Flores et al., 2010; Kagohara, Sigafoos, Achmadi, O’Reilly, & Lancioni, 2012), iPods (e.g., Gihak, Fahrenkrog, Ayres, & Smith, 2010; Kagohara et al., 2011), and mobile phones (e.g., Bryen, Carey, Friedman, & Taylor, 2007; McMahon, Cihak, Gibbons, Fussell, & Mathison, 2013; Taber, Alberto, Seltzer, & Hughes, 2003).

While they are used more and more frequently to teach new skills, computers and mobile technologies are most commonly used to target a specific skill deficit. In their review of the literature regarding the use of technology and mobile devices (e.g., iPad, iPods, mobile phones) for individuals with developmental disabilities, Kagohara et al. (2013) came to several conclusions. Most notably, they concluded that while technology is a viable aid for individuals with developmental disabilities, they are used primarily only as intervention delivery systems (e.g., presenting instructional video, prompting students). Further, research has indicated that students with ID often do not take full advantage of technology (Happiestad, 2007; Kling & Wilcox, 2010; Tanis et al., 2012), and that they are less likely to have access to and benefit from technology when compared to students with other disabilities (Wemeyer, Smith, Palmer, & Davis, 2004).

The concept of digital literacy was first introduced in 1997, described as “the ability to understand and use information in multiple formats from a wide range of sources when it is presented via computers” (Gilster, 1997, p.1). Many variations in the definition of digital literacy have since emerged, with job-specific applications ranging from the technical aspects of operating in digital environments to the cognitive and socio-emotional aspects of a computer-driven environment (Eshet, 2004). The educational literature uses the term “21st century literacy” to describe this phenomena, applying to “the ability to read and interpret media (text, sound, images), to reproduce data and images through digital
manipulation, and to evaluate and apply new knowledge gained from digital environments” (Jones-Kavalier & Flannigan, 2006, p. 9). Incorporating the current technologies used today with the prerequisite skills potentially required for future technology, 21st century literacies focus on preparing students with the skills needed to participate successfully in tomorrow’s workplace and classroom (Collier, 2007). The shift from a print-based to a screen-based society calls for teachers to integrate digital literacy skills into instruction to prepare students for expectations of potential employers; therefore, digital literacy is considered to be an essential life skill (Bawden, 2008).

Digital literacy skills have become an essential part of learning how to function independently in today’s society. In addition to the use of technology for traditional instructional purposes (e.g. computer-assisted instruction) to promote academic progress and outcomes, students with disabilities can benefit from technology to support learning in the area of life skills (Wehmeyer, Smith, Palmer, & Davies, 2004). However, a 2005 study surveyed 83 adults with ID regarding their use of various technologies, and found that only 41% of these respondents used a computer, 25% used the Internet, and 11% used electronic organizers; statistics all well below the general population’s use of such devices and technologies (Carey, Friedman, & Bryen, 2005). Functional use of common internet activities for communication and information pursuits require a level of functional literacy and a range of cognitive skills including decoding, comprehension, and written expression (Johnson, 2007). A lack of functional digital literacy, the basic skills required for users to operate effectively in digital environments, could soon be the only factor preventing access for all. It is imperative for students with ID to learn functional digital literacy skills to be fully included in educational settings, to increase their access to digital materials, and to access employment opportunities.

Frequently cited as the most common online activity (U.S. Census Bureau, 2010), email has become the dominant form of communication in workplace and educational settings. A working email account is also required for access to many online social communications and networks (e.g. Twitter, Facebook). Multiple cognitive processes are involved when a user engages in online communication in the form of email and other text-based digital messaging services: basic literacy to accurately input intended email address(es), comprehension to create or respond to a message, discrimination between relevant and irrelevant information, and interpretation of the email sender’s intent (Johnson, 2007). Interpretation of digital messages requires a different skill set than face-to-face communication, where additional information can be obtained through the speaker’s facial expressions and use of gestures (Szewczak & Snodgrass, 2002). In addition, email addresses are becoming increasingly associated with one’s online identity. A working email address is necessary to access educational information and communications, complete financial activities (e.g. banking and online purchasing), participate in social networking, and even access newsletters and job opportunities. Yet, in a recent study, researchers found that only 18.7% of people with ID have an email address (Palmer, Wehmeyer, Davies, & Stock, 2012).

A new trend in online media is the use of social bookmarking tools, such as Pinterest, Reddit, and Diigo, which allow internet users to “bookmark” websites for later use. Users can visually scan for bookmarks they wish to share with other users or social media sites. Frequently used by academic libraries, social bookmarking is a practical tool that can be used to share leisure interests and increase access to employment opportunities through social networking and social media websites (Redden, 2010). In 2011, The National Council on Disability (NCD) found that social networking may be more critical for people with disabilities than for the general population (NCD, 2011). Social bookmarking is an age-appropriate life skill that can be used for educational, employment, or leisure purposes, and therefore, is considered to be a functional skill.

Cloud storage is another functional tool that provides access and availability to users.
The development of cloud storage services has profoundly impacted the availability and accessibility of digital information. Any document, photo, video clip, or other file type uploaded to the cloud may be accessed or shared anywhere on any computer or device using an internet connection. For students with ID, the ability to access files stored in the cloud allows opportunities for increased independence and the ability to self-manage various supports as needed. These supports include communication notebooks, vocational routines and checklists, picture prompts and picture schedules, video models, social stories, and safety skills. Rather than carrying cumbersome devices and supportive tools, users now have digital access to the same supports. Students with disabilities can use a variety of mainstream technologies and readily available devices to access these supports. Many organizations, including educational institutions, are investing in cloud storage as an effective, and cost efficient means of providing access to digital materials, increasingly becoming the only option for user access. Recently, four of the top five essential apps for college students with ID included cloud storage of files for easy retrieval and sharing (Cooney, 2012).

Despite the increased importance of functional digital literacy, the current research is limited on teaching specific functional digital literacy skills to students with ID. Specifically, no research has been conducted on teaching high school students with ID functional digital literacy skills, namely: (a) receiving and sending emails, (b) managing social bookmarks, and (c) accessing documents through a cloud storage service. The purpose of this study is to examine the effect of using systematic digital literacy instruction to teach these three functional digital literacy skills to high school students with mild and moderate ID. In addition, students’ opinions of digital literacy skills were surveyed following the study.

Method

Participants and Setting

Three high school students participated in this study: one male student (Charles) and two female students (Kate and Faith). Participants were selected based on the following:

- (a) diagnosis of an intellectual disability, (b) no email address, (c) no physical disability which impeded the performance of the skill, and (d) agreeing to participate in the study.

Table 1 details characteristics for each student. Charles also had a comorbid diagnosis of autism. In addition, all students participated in a life skills curriculum, which included daily living skills, community skills, and career goals. Further, five students without disabilities worked in the classroom as peer tutors. All phases of this study occurred in the students’ special education resource classroom. Students also worked toward communication and social skills goals included in their Individualized Education Plans (IEPs).

The setting was a high school special education classroom, which included a special education teacher and teacher assistant. The special education teacher had 10 years of experience teaching students with intellectual disabilities, and the teacher assistant had two. The special education teacher and teacher assistant have worked together for the past two years. The special education teacher implemented all intervention procedures.

Materials

This study incorporated the use of three Dell Vostro 270 desktop classroom computers with Windows 7 operating systems. We created Gmail accounts for email, Diigo accounts for social bookmarking, and DropBox accounts for cloud storage for each of the three participating students. Each of these digital tools are available free of charge. Once accounts were created, the teacher instructed the students to
select a password they could easily remember. Students also were provided initially with an index card with their username and password printed in a 14-point font for reference if needed. The index card was made available when the student entered his or her username and password incorrectly. Since these tasks required a username and password, we decided to provide non-contingent access during baseline to increase the likelihood of students performing later skill steps independently. During intervention, students could request the index card if needed during the intervention phase. As students consistently entered their username and password independently (i.e., 3 consecutive sessions), the index card was stored in a locket cabinet for security purposes although students could still request the card at anytime. Additionally, screenshots of each task-analyzed step for all three target skills were created and printed. Screenshots of each motor demonstration step were used as visual prompts and incorporated during general digital literacy instruction.

Variables and Data Collection

Event recording procedures were used to record the number of digital literacy task-analyzed steps completed independently or with teacher assistance. The dependent variable was the number of task-analyzed steps completed independently, which was calculated as the percentage of task analysis steps completed independently, to email, bookmark, and use cloud storage independently. Table 2 lists the task-analyzed steps for each targeted skill. For emailing, students were expected to (a) sign-in to their email account, (b) respond to an email sent by their teacher, and (c) compose and send a new email to a fellow student (i.e., peer tutor). A total of 13 task-analyzed steps were required to independently email. For bookmarking, students were expected to (a) save/bookmark at least one career related website, and (b) access at least one saved website/bookmark. A total of 10 task-analyzed steps were required to independently use bookmarking. For cloud storage, students were expected to (a) sign-in to cloud storage, (b) download a document, (c) revise the document, and (d) upload a document. A total of 17 task-analyzed steps were required to independently use cloud storage.

An independent response was defined as initiating the first step in the task analysis within 10 s and completing each step within 30 s without teacher assistance, except for the
first step of social bookmarking (i.e., find page to bookmark). Although students were required to initiate the task within 10 s, they had 5 min to review and find at least one job related website to bookmark. The additional 5 min was provided only for step one; students were expected to complete the remaining steps (i.e., 2–10) within 30 s. If a student did not independently initiate the first step within 10 s, complete a step within 30 s of initiation, or 5 min to complete step one for social bookmarking, teacher assistances were implemented using a system of least prompts (Ault & Griffen, 2013). The least-to-most prompt hierarchy consisted of the following levels: (a) verbal prompt, (b) gesture, (c) gesture plus verbal explanation, (d) modeling plus verbal explanation, and (e) physical assistance plus verbal explanation. The percentage of steps completed independently was graphed for visual analysis. The special education teacher recorded the data and implemented all instructional procedures for all sessions, and one session occurred per day.

Maintenance data were collected 9 weeks later after the student reach acquisition criteria. Similarly, event recording procedures were used to record the number of digital literacy task-analyzed steps completed independently or with teacher assistance and then calculated as the percentage of task analysis steps completed independently for each student. During the 9 weeks, students continued to use each skill daily.

**Design**

A multiple probe design across behaviors (Gast & Ledford, 2010) was used to examine a possible relation between digital literacy instruction and the acquisition and maintenance of three digital literacy skills. All digital literacy skills were probed during baseline prior to instruction. After baseline data were stable, the teacher implemented general digital literacy instruction and targeted emailing, while social bookmarking, and cloud storage skills continued to be probed during the baseline phase. Contingent upon the student reaching the acquisition criterion for emailing (i.e., 100% for three consecutive sessions), the teacher implemented digital literacy instruction and targeted social bookmarking, while cloud storage continued to be probed during the baseline phase. Lastly, the teacher implemented digital literacy instruction and targeted cloud storage after the student reached criterion for bookmarking. Acquisition criteria were defined as performing all task-analyzed steps of the targeted digital literacy skill independently for three consecutive sessions.

**Procedure**

**Baseline.** Materials for each task were pre-arranged and readily available to the student. The teacher began all sessions by asking the student to sit at the computer and either “check your email, research career websites, or check your schedule.” Students were given 10 s to initiate the first step in the task analysis and 30s to complete each step, except for step one of bookmarking in which students were provided 5 min to review and save/bookmark a website. In addition, no tangible reinforcement was given. The baseline phase continued for a minimum of three sessions or until stability was achieved.

However, the teacher assisted with spelling when students requested and provided the student’s username and password on an index card. Since all skills required an email address and password, all students were registered for an email account accessible through a username and password. The students self-selected a username and password. Students were provided with an index card with their username and password printed in a 14-point font and instructed to refer to the card if they forgot their login information. The students’ username and password was provided to increase the likelihood of students being able to demonstrate the most digital literacy skills possible. Since a username and password was required to progress to other steps of the skills, we did not want to artificially suppress a student’s performance of specific skills that they may otherwise be able to demonstrate.

For email, students were expected to sign-in, and respond to a daily email from their teacher, which included one closed-ended question (e.g., are you going to the community to work today, who won the game, what are you eating for lunch, what’s the weather like). In addition, students were expected to compose an email to a peer tutor. Students
were encouraged to email the peer tutor about what they were doing today, what they did over weekend, or plans for the upcoming weekend. Peer tutors’ email addresses were previously saved as contacts; therefore, as soon as the student began typing the first letters of a person’s name, their corresponding email address appeared.

For bookmarking, students used a common social bookmarking application (i.e., Diigo) and search a statewide jobs website, which provides any job seeker with a list of open positions within a predetermined distance (e.g., 5, 10, 25, 50 miles) of an entered zip code. The students’ zip code and a 10-mile radius was preset for all students. Students were expected to review open employment positions, bookmark at least one career website of interest and to access at least one bookmark website.

For cloud storage, students were expected to download an activity posted by their teacher from a common cloud storage website, complete the activity, and then upload the completed work. Students were expected to download a daily schedule task checklist. The checklist included pictures of tasks completed the previous school day and the current day’s schedule of tasks that required completion. Each student was expected to download the schedule, self-report by typing a “Y” next to each task indicating they completed it, review the current daily schedule, and upload the new schedule or document to the cloud storage.

Digital Literacy Instruction

General instruction. After the baseline data stabilized, the special education teacher reviewed each screenshot and verbally described each task-analyzed step three times. Then, the teacher demonstrated and verbally explained each task-analyzed step three times. This review and demonstration occurred during one session and lasted approximately 30 min. On the following school day, the teacher implemented a system of least prompts. First, general digital literacy instruction targeted emailing. Contingent on skill acquisition, general literacy instruction was systematically implemented for the second skill (i.e., bookmarking) and then the third skill (i.e., cloud storage).

Emailing. On the following school day, the teacher cued the students to “check your email.” Students were given 10 s to initiate the first step in the task analysis and 30 s to complete each step after being cued to “check your email.” If they did not independently complete a step within 30 s of initiation, teacher assistance was provided. The teacher administered a system of least prompts (Ault & Griffen, 2013). The least-to-most prompt hierarchy consisted of the following levels, which included a 4-s level between each prompt level: (a) verbal prompt (e.g., “do you see where you need to login?”), (b) gesture (e.g., pointing to login box where to login), (c) gesture plus verbal explanation (e.g., pointing to the login box and providing a verbal explanation), (d) modeling plus verbal explanation (e.g., pointing to the login box, providing a verbal explanation, and demonstrating the correct response), and (e) physical assistance plus verbal explanation (e.g., holding the student’s wrist, pointing to the picture prompt, providing a verbal explanation, and physically assisting the student the correct response). Digital literacy instruction continued until the student reached acquisition criteria (100% independence for three consecutive sessions) for emailing including signing-in, responding to an email, and composing and sending an email.

Bookmarking. Following the acquisition of emailing, the teacher implemented one session of the general digital literacy procedures. Afterwards, the teacher cued the student to “research career websites.” Similar to emailing, students were given 10 s to initiate the first step in the task analysis and 30 s to complete each step of the bookmarking task analysis, except for step one (i.e., find page to bookmark). Students were given 10 s to initiate the skill and 5 min to complete this skill step; students were expected to complete the remaining steps (i.e., 2–10) within 30 s. If they did not independently complete a step within 30 s of initiation or 5 min for step one, teacher assistance was provided and the system of least prompts was implemented. Instruction continued until the student reached acquisition criteria of saving and accessing at least one
website (i.e., bookmarking) with 100% independence for three consecutive sessions.

Cloud Storage. Following the acquisition of bookmarking, the teacher implemented the general digital literacy procedures for one school day. Afterwards, the teacher cued the student to “check their schedule.” Similar to previous digital literacy skills, students were given 10 s to initiate the first step in the task analysis and 30 s to complete each step. If they did not independently complete a step within 30 s of initiation, teacher assistance was provided and the system of least prompts was implemented. Instruction continued until the student reach acquisition criteria of signing-in, downloading, and uploading a document (i.e., cloud storage) with 100% independence for three consecutive sessions.

Social Validity

Following the completion of the study, students completed a social validity questionnaire to assess their opinions regarding the acquisition of digital literacy skills. The research assistant administered the questionnaire to each student individually. It consisted of a 14-item Likert scale, ranging from 1 (indicating strongly disagree) to 5 (strongly agree). The possible scores ranged from 14 to 70. Higher scores indicate greater social acceptance. The research assistant also read each item and response options aloud. Following the 14 items, an open-ended question asked students, “What else do you think?”

Interobserver agreement (IOA) and Treatment Fidelity

The classroom teacher and a graduate assistant independently and simultaneously collected interobserver agreement (IOA). Interobserver agreement data were collected during a minimum of 50% of baseline, intervention, and all maintenance sessions for each student. Observers independently and simultaneously recorded the number of steps each student completed independently. In addition, the peer tutors were instructed to forward the email to the teacher and graduate assistant that was composed and sent by the participating students. Interobserver agreement was calculated by dividing the number of agreements of each student’s responses by the number of agreements plus disagreements and multiplying by 100. Interobserver agreement was 100% for each participant’s assessment during all phases.

Two research assistants also collected treatment fidelity data during a minimum of 50% intervention and all maintenance sessions for each student. The teacher was required to provide access to the computer and instruct students to email, social bookmark, or use cloud storage. The teacher was observed providing general digital literacy instruction including reviewing screen captions and demonstrating each skill while providing verbal explanations. In addition, the teacher implemented the system of least prompts with a 4-second delay between hierarchical levels. The research assistants calculated treatment fidelity levels by dividing the number of observed teacher’s behaviors by the number of planned teacher behaviors and multiplying by 100. The mean treatment fidelity was 88% for Kate, 95% for Charles, and 98% for Faith. The majority of disagreements occurred as a result of not waiting a full 4 seconds before introducing the next level of prompt.

Results

Overall, the students readily acquired each digital literacy skill. On average, students completed few task analysis steps independently during baseline. During digital literacy email instruction, all students independently replied to the teacher’s emails, indicating they completed the steps of the task analysis, as well as comprehended the message received. In addition, all students composed and sent a new email to a peer. During digital literacy instruction, students required a mean of 11 sessions to reach 100% independence for 3 consecutive sessions for replying and sending an email (M = 73%), 9 sessions to reach criteria for social bookmarking (M = 77%), and 13 sessions to reach criteria using cloud storage (M = 71%). Each student demonstrated an immediate increase in learning to email and social bookmark with a 100% nonoverlapping data. Kate also demonstrated an immediate increase in learning to use cloud storage with a 100% nonoverlapping data; Faith and Charles’ percentage of nonoverlap-
ping data for cloud storage was 87.5% and 92%, respectfully. In addition, all students continued to use the digital literacy skills daily and maintained each digital literacy skill nine weeks later.

Kate. Figure 1 displays the percentage of task-analyzed steps Kate completed independently for emailing, social bookmarking, and using cloud storage. During baseline, Kate did not complete any of the task analysis steps for emailing or using social bookmarking. She did complete the first two steps for cloud computing independently (i.e., entering username and password). During digital literacy instruction, Kate required 10 sessions to acquire emailing independently for 3 consecu-
tive sessions \((M = 75\%, \text{ range } = 23–100\%)\), 7 sessions to acquire social bookmarking independently for 3 consecutive sessions \((M = 76\%, \text{ range } = 30–100\%)\), and 9 sessions to acquire cloud storage independently for 3 consecutive sessions \((M = 73\%, \text{ range } = 25–100\%)\). Across all skills, she immediately demonstrated an ascending trend and 100% nonoverlapping data. Kate also maintained each digital literacy skill nine weeks later.

*Faith.* The percentage of task-analyzed steps Faith completed independently for emailing, social bookmarking, and using cloud storage is displayed in Figure 2. During baseline, Faith performed similarly to Kate. Faith did not complete any of task analysis steps for emailing and using social bookmarking, but independently completed the first two steps of cloud storage (i.e., entering username and password). She required 15 sessions to acquire emailing with 100% independence for 3 consecutive sessions \((M = 66\%, \text{ range } = 12–100\%)\), 10 sessions to acquire social bookmarking independently \((M = 71\%, \text{ range } = 30–100\%)\), and 16 sessions to acquire cloud storage independently \((M = 67\%, \text{ range } = 19–100\%)\). Faith immediately began learning how to email and use social bookmarking as demonstrated by 100% nonoverlapping data. Although a general ascending trend was observed for cloud storage, her skill acquisition was more gradual with a demonstration of 87.5% nonoverlapping data. In addition, Faith maintained each skill nine weeks later.

*Charles.* Figure 3 displays the percentage of task-analyzed steps Charles completed independently for emailing, social bookmarking, and using cloud storage. During baseline, Charles independently completed the first two steps independently (i.e., entering username and password) for emailing and using cloud storage. He did not complete any of the task analysis steps for using social bookmarking during baseline. During digital literacy instruction, Charles required 8 sessions to email with 100% independence for 3 consecutive sessions \((M = 79\%, \text{ range } = 34–100\%)\), 10 sessions to independently use social bookmarking \((M = 83\%, \text{ range } = 30–100\%)\), and 13 sessions to independently use cloud storage \((M = 72\%, \text{ range } = 19–100\%)\). Charles immediately demonstrated an ascending trend and 100% nonoverlapping data for learning to email and use social bookmarking. Charles also readily learned to use cloud storage with a demonstration of an ascending trend and 92% nonoverlapping data. In addition, Charles maintained each skill nine weeks later.

**Social Validity**

Table 3 displays each student’s rating on the 14-item social validity questionnaire. The questionnaire scores could range from 14 to 70 with a higher score indicating greater social acceptance. Kate reported a total score of 66, Faith reported a total score of 63, and Charles reported a total score of 60. Students indicated that they liked all three skills. Emailing and social bookmarking were most preferred by the students, as they indicated that it would be useful to communicate with friends, family, co-workers, and employers. They also indicated that they would continue to use both email and social bookmarking, and would recommend social bookmarking to other students. Kate suggested that she really liked how she can easily find websites about her favorite TV shows. Students were indifferent about the use of cloud computing. Although Kate and Faith indicated some future use of the skill, Charles indicated that he would not continue using cloud storage.

**Discussion**

The purpose of this study was to examine the effects of teaching individuals with moderate intellectual disabilities to use email, social bookmarking, and cloud storage. All students readily acquired each of these digital literacy skills. All students also readily acquired the skill of entering their username and password, however, they were unfamiliar with most of the other steps to complete each skill independently. With the implementation of digital literacy instruction including the teacher’s review of static pictorial screen shots, modeling each digital literacy skill, and proving teacher assistance in the form of the system of least prompts, all students acquired and maintained the functional digital literacy skills. A functional relation was established; experimental control was demonstrated in at least three different series at three different points in time for systematic digital literacy instruc-
tion and emailing, social bookmarking, and using cloud storage independently.

These findings support previous findings that suggest that systematic instruction is an effective strategy to teach new skills (Spooner et al., 2012), computing skills (Pennington, Stenhoff, Gibson, & Ballou, 2012) and digital media skills (Cihak, Fahrenkrog et al., 2010) to learners with intellectual and developmental disabilities. This study extends the

Figure 2. Faith’s percentage of steps completed independently across digital literacy skills.
existing body of research by demonstrating the use of systematic instruction for teaching digital literacy skills to students with mild to moderate ID.

This study extended the current literature in several ways. First, by demonstrating the effectiveness of teaching functional digital literacy skills, it also demonstrates the ability of such instruction to provide students greater access to the digital society. Burgs-

Figure 3. Charles’ percentage of steps completed independently across digital literacy skills.
Tahler (2002) noted a digital divide exists between people with disabilities and people without disabilities. Digital literacy is a critical life skill in today’s digital society. Instead of utilizing such technology to only target specific skill deficits and/or as intervention delivery systems (i.e., Kagohara et al., 2013), it is important to employ such technology in order to teach students with ID how to access digital communities. When considering the rise in availability of new software and applications, teaching students with ID the fundamentals of such technology will provide them greater access to and use of the new technology. Rather than using technology as a cul-de-sac, acquiring and using digital literacy skills allows users to access an on-ramp of information and social contacts for employment, interactions, and independent living. Digital technology can be a resource for learners with ID in personal, academic, and employment pursuits.

An additional benefit of the intervention was the expansion of language through email. Early in the intervention, students simply responded to the teacher’s email with one word or short phrases, but as the intervention progressed, they sent more elaborate statements and details. Additionally, after Kate acquired the first skill, she emailed the researcher from her home computer without instruction from the researcher. This indicates that Kate generalized the skill to another setting and proves email to be a socially valid skill for Kate.

Second, this study incorporated the use of essential digital tools (i.e., email, social bookmarking, and cloud computing). Because digital literacy skills are necessary for participation in education, the gap is widened for those without the skills. Researchers indicated that 85% of people without disabilities use a computer or other device to access the Internet in comparison to only 54% of those with disabilities (NCD, 2011). The National Center on Educational Statistics (NCES) reported that children aged 5 to 17 years without a disability were significantly more likely to use computers and the Internet than their peers with disabilities (2001). Children and adolescents with ID were even less likely to use a computer or the Internet. Technology can either stand as an obstacle or a facilitator in education for individuals with disabilities (Fichten et al., 2001). Teaching students to use digital tools such as email, social bookmarking, and cloud storage enables them to be more fully included as digital citizens.

### Table 3

<table>
<thead>
<tr>
<th>Items</th>
<th>Kate</th>
<th>Faith</th>
<th>Charles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning to email is a useful skill</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Learning how to email will help me communicate with friends and family</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Learning how to email will help me communicate with supervisors and co-workers</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>I will continue to use email</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>I liked learning how to use email</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Learning to use social bookmarking is a useful skill</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Social bookmarking will help retrieve my important websites</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>I will continue to use social bookmarking</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>I will tell other students about social bookmarking</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>I liked bookmarking sites I visit often</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Learning to use cloud storage is a useful skill</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>I will continue to use cloud storage</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>I liked learning how to use cloud storage</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>I will tell other students about cloud storage</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>66</td>
<td>63</td>
<td>60</td>
</tr>
</tbody>
</table>

*Note. Likert scale was strongly disagree (1), disagree (2), neither agree or disagree (3), agree (4) and strongly agree (5).*
From elementary schools to college campuses, the use of technology is widespread. Mobile devices and electronically formatted textbooks are now commonly found in the classroom. The portability of such technology enables users with disabilities to access supports from anywhere in the world. Equipped with digital skills, people with intellectual disabilities can communicate with friends and family members via email, bookmark employment websites, and retrieve saved documents via cloud computing.

Third, the current study provided supports to facilitate work readiness skills. Digital communication is vital to establish connections with friends, communicate with coworkers, and network with other professionals. In this study, students were instructed to respond to emails, bookmark websites of interest, and store documents in the cloud for later use. The overarching goal of secondary education for students with ID is to prepare them to be productive, wage-earning citizens. To prepare for the competitive workforce, learners with ID must be equipped with digital skills such as using username and passwords, email usage, and internet access. As a component of Information and Communication Technology (ICT), digital communication is an important skill offering individuals with ID a means of communicating with others for work and/or leisure purposes. It can ease social isolation and advance academic, career, and leisure goals by connecting people with ID to a community of peers and a network of supports (Burgstahler, 2002). People with ID can use email to communicate and socialize with friends, family members, teachers, and employers. Email is the primary means of communication outside of school between employers and workers. Email has become an environment for conducting work, and for maintaining social life.

Participants of this study ranged in levels of ability, and all students were successful in acquiring the digital literacy skills. Students also indicated that the skills were useful and enjoyable, particularly emailing. These skills allow high school students the opportunity to communicate through email, to save important websites, and to obtain information and supports to complete academic and non-academic tasks. These are all necessary skills for students preparing to transition into life after school. They facilitate the job application process and can be used to increase numerous other educational skills.

Although this study indicated positive outcomes, conclusions must be interpreted within the context of this study and several limitations should be considered. The participating students in this study were all diagnosed with an intellectual disability, which limits generalizability to other populations. Further investigations that include participants with a wide range of characteristics are warranted.

A second limiting factor was response generalization. As the students were learning to email, they generalized entering their username and password to access and use cloud storage. Due to the number of shared steps of the task analysis, future studies can address this limitation by reducing the number of common steps among specific digital literacy skills. Third, the technology itself is a third limiting factor. Many students may be unfamiliar with digital tools such as social bookmarking and cloud computing. This would prevent the accessibility and use of these skills. Also, these applications require the use of internet access, preferably high speed internet access.

As the findings from the study suggest, incorporating functional digital literacy skills into curriculum for secondary students with intellectual disabilities may be effective at improving functional, academic, and independent living skills. Future research should include the establishment of a framework of recommended practices for educators to provide effective systematic instruction of these critical skills. Additional target skills should include instruction for using internet search engines, utilizing built-in accessibility features, and the use of mobile applications. Students also were provided an index card with their username and password until they could consistently enter their security information independently. A more advised security practice that assists students to create an easy to remember but challenging password is needed.

Although teaching digital literacy skills to students with mild to moderate ID is an ongoing endeavor for educators and requires sustained effort, this study demonstrates positive outcomes for the participating students. A stu-
dent’s disability is not the defining characteristic of the participant, nor is it necessarily a barrier to digital literacy participation. The Internet can be a wonderful resource for students with ID. They can use it for school, communicating with others, searching for job opportunities, building a network of supports, and playing interactive games. Students who are old enough to type in a few letters on the keyboard can literally access the world. The first step in the digital literacy process is ensuring that students with ID can demonstrate various essential digital skills. This study demonstrated the feasibility of increasing digital literacy by teaching students how to use these skills effectively and over time. With the advancement of technology, accessibility features, and user-friendly design, people with ID can participate as active members of a digital society.

References


Davies, D. K., Stock, S., & Wehmeyer, M. L. (2002b). Enhancing independent time-management skills of individuals with mental retardation using a


Received: 30 January 2014
Initial Acceptance: 1 April 2014
Final Acceptance: 26 July 2014
Review of Evidence-based Mathematics Interventions for Students with Autism Spectrum Disorders

Juliet E. Hart Barnett and Shannon Cleary
Arizona State University

Abstract: Students with autism spectrum disorders (ASD) are being included more frequently in the general educational setting, and are therefore increasingly expected to access and master core curricular content, including mathematics. However, mathematics often presents challenges to students with ASD. Interventions to improve the mathematics skills of students with ASD have been recommended. This comprehensive literature review synthesized eleven studies of mathematics intervention strategies for students with ASD. Though studies related to instructional interventions in mathematics for students with ASD are limited, these students can benefit from mathematics interventions, which can help them strengthen their mathematics skills, increase independence when completing problems, and use acquired skills in community or other applied settings. Future implications include the need for additional, empirically-supported interventions in mathematics for students with ASD and the need to target more academically-oriented math interventions for this population, particularly in the context of problem solving, which will assist in determining the potential of students with ASD to achieve mathematical success.

The education of students with autism spectrum disorders (ASD) in inclusive classroom environments is becoming increasingly common (Cihak, Fahrenkrog, Ayres, & Smith, 2010). Placement of students on the spectrum in general education has increased more quickly than all other disability categories combined (Sansoti & Powell-Smith, 2008). Today, 36 percent of students with ASD spend more than 80 percent of their school day in general education classes, which constitutes a significant increase from their 4.8 percent inclusion rate in 1991 (Whitby, 2013). As a result, although students’ individual IEP goals and needs determine special education services, there is a rising expectation that these learners will access and master the same core curricular content as their typically developing peers and ushers in a related demand for effective educational interventions to promote students’ successful content acquisition (Knight, Smith, Spooner, & Browder, 2012).

However, instructional programs for students with ASD tend to focus on communication and social skills (Plavnik & Ferreri, 2011; Wang & Spillane, 2009) as well as functional and life skills as opposed to traditional age and grade level content areas (Cihak & Grim, 2008; Rayner, 2011). Within the research base related to academic content, the focus for students with ASD is primarily related to reading (Bouck, Satsangi, Taber-Doughty, & Courtney, 2014; Delano, 2007). Although mathematics education is a national priority for all students (Ellis & Berry, 2005), there are fewer comparable, empirically supported interventions in mathematics than in literacy for students with ASD (Bouck et al., 2014).

Mathematics is an area of academic concern for students with ASD. Nearly 25 percent of students with ASD contend with a mathematics learning disability (Mayes & Calhoun, 2006). A recent longitudinal study of individuals with ASD indicated slower growth rates in calculation skills as compared to students with learning disabilities (Wei, Lenz, & Blackorby, 2013). In addition, although many of these learners maintain adequate mathematics performance in the earlier grades when rote memorization of facts and procedures is im-
portant (Chiang & Lin, 2007), the same students may struggle as they enter middle school when the content becomes more abstract, cognitively complex, and emphasizes problem solving, higher level thinking, and mathematical reasoning, which are cited by researchers as areas of weakness for children with ASD (Mayes & Calhoun, 2003; Whitby & Mancil, 2009).

The difficulties students with ASD encounter in mathematics may likely stem from deficits in executive functioning, including planning, organization, working memory, mental flexibility, attention, self-monitoring, and impulse control (Happe, Booth, Charlton, & Hughes, 2006; Hughes, Russell, & Robbins, 1994). Additionally, the language impairment associated with ASD may also negatively impact mathematics development across several areas such as number-word sequence, calculation, fact retrieval, and in particular, problem-solving (Donlan, 2007; Zentall, 2007), which requires students to utilize both semantic and numeric information (Rockwell, Griffin, & Jones, 2011). Despite the fact that researchers have investigated the role of these student-level characteristics on mathematical skills, they also recognize the difficulty in extrapolating the extent to which these challenges impact math performance given the limited literature base (Chiang & Lin, 2007; Fleury et al., 2014), which could also suggest that ASD students’ difficulties with mathematics may be related to a lack of targeted direct instructional opportunities in this area. Notwithstanding, the general education curriculum and state assessments in mathematics increasingly emphasize the importance of developing the conceptual understanding and problem solving described and recommended by the National Council for Teachers of Mathematics (NCTM) across domain areas (Botte, 2001; NCTM, 2002; Rockwell et al., 2013), which students with ASD who are educated in inclusive settings will likely encounter.

Given the rise of students with ASD in inclusive settings, their documented difficulties in mathematical understanding and problem solving, and the prominence the NCTM has placed on conceptual understanding and problem solving across skill areas, examining and evaluating the extant research in the area of mathematics interventions for learners with ASD is a national research priority. The purpose of this review is to contribute to the current knowledge base on effective mathematics interventions for students with ASD in order to ensure that their teachers consider and implement effective teaching techniques to help their students acquire the mathematical skills identified on their IEPs and those being taught in the general education classroom, and achieve maximum success.

Method

To be included in this comprehensive review, published articles had to meet the following criteria: (a) be published in a peer-reviewed journal, (b) include students of any age ranging from elementary age through post secondary education, (c) include participants identified with an ASD diagnosis (i.e., autism, PDD-NOS, or Asperger syndrome), (d) be conducted in any educational or outside clinical, home-based, or tutorial setting (e.g., general education classroom, special education resource or self-contained classroom, etc.), and (e) describe research studies evaluating the effectiveness of academic or functional interventions targeting mathematical content standards (i.e., number and operations, algebra, geometry, measurement, and data analysis/probability) and/or process standards (i.e., problem-solving, reasoning, communication, connections, and representations) as described by the NCTM.

During the review process, exclusionary criteria were also developed. First, studies lacking an empirical research design were excluded. That is, articles featuring mathematics teaching strategies or tips for students with ASD were eliminated to include only articles featuring an administered intervention strategy supported with data/evidence to determine the effectiveness of that intervention strategy. Second, articles that featured mathematics problems as a component of a larger investigation of a teaching strategy, but did not use the study of mathematics interventions as the primary goal of the experiment, were excluded from the literature review.

The search was conducted electronically using online ERIC and Academic Search Premier databases with no limits placed on year of publication. Key search terms pertained to autism and mathematics (e.g., autism spec-
trum disorder, Asperger’s Syndrome, autism, intellectual disabilities, mathematics, mathematics instruction, teaching strategies, interventions, etc.) From this initial search, journals typically publishing intervention research in ASD were identified: Journal of Autism and Developmental Disorders, Focus on Autism and Other Developmental Disabilities, Research in Autism Spectrum Disorders, Behavioral Interventions, Preventing School Failure, and Education and Training in Autism and Developmental Disabilities. Furthermore, multiple mathematics journals were also searched for any additional articles that met the search criteria. Mathematics education journals included in the search were: Journal for Research in Mathematics Education, Journal of Mathematics Teacher Education, Educational Studies in Mathematics, and Mathematics Education Research Journal. These journals were electronically hand searched for all abstracts of interventions related to improving the mathematics knowledge or skills of learners with ASD.

To ensure reliability of the search process, the two authors of the study and a graduate student independently conducted the initial review. Reliability was then calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. Inter-observer agreement (IOA) was 85%. Through closer examination of all studies initially located as well as discussion of inclusionary and exclusionary criteria, 100% IOA was achieved. Similar reliability procedures were conducted to ensure consistency across observers in the analysis of each study, and 100% IOA was achieved through consensus. In sum, the search was conducted to obtain and review all available research related to mathematics education interventions that have been conducted with students with ASD and that had as their primary goal increasing these students’ mathematical skills in alignment with NCTM standards.

Results
The initial search produced 13 studies, of which, two were eliminated based on failure to meet inclusionary criteria. Therefore, eleven studies met the aforementioned inclusionary criteria. Table 1 shows the participants, setting, intervention, intervention testing, design, dependent variables, independent variables, and treatment effects/results of each study reviewed. Studies were largely single-subject design, were grouped by similarity of intervention type, and were designated as an intervention implementing either visual representations or some form of cognitive strategy instruction to target specific mathematics skills. In addition, studies will be described in terms intervention setting (i.e., self-contained vs. general education classroom) as well as the mathematics skill targeted (i.e., academic vs. functional skills). Lastly, implications for practitioners and directions for future research will also be discussed.

Of the 11 studies, six implemented visual representations to develop specific math skills (Bouck et al., 2014; Burton, Anderson, Prater, & Dyches, 2013; Cihak & Foust, 2008; Fletcher, Boon, & Cihak, 2010; Rockwell et al., 2011; Waters & Boon, 2011). Visual representations include manipulatives, pictures, number lines, and graphs of abstract concepts, functions, and relationships (Xin & Jitendra, 1999). Representation approaches to solving mathematical problems include pictorial (e.g., diagramming), concrete (e.g., manipulatives, or objects that assist students in understanding abstract mathematical concepts or properties by representing them in multiple ways), verbal (linguistic training), and mapping instruction (schema-based) (Xin & Jitendra, 1999). The 5 remaining studies implemented some form of cognitive or meta-cognitive strategy instruction (Banda & Kubina Jr., 2010; Cihak & Grim, 2008; Hua, Morgan, Kaldenberg, & Goo, 2012; Rapp et al., 2012; Whitby, 2013). Cognitive strategies are frequently used in mathematics problem solving and are defined as a series of sequenced procedures that permit a student to complete a task effectively using rules, processes, and steps that are applied systematically to obtain a problem solution (Simpson, 2005), and include the meta-cognitive elements of when and where to apply specific strategies in the implementation and evaluation of the process and outcome (Montague, 2008; Reid & Leinemann, 2006). Both of these approaches are among those considered evidence-based for improving mathematical knowledge and skills of students with disabilities, particularly those who are low achieving and those with learning disabilities (Baker, Gersten, & Lee, 2002), and
## TABLE 1

** Reviewed Studies by Instructional Intervention Type **

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants in Experiment (Total Number of Participants, Number of Participants with ASD, Ages)</th>
<th>Setting</th>
<th>Intervention</th>
<th>Intervention Testing</th>
<th>Design</th>
<th>Dependent Variables</th>
<th>Independent Variables</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bouck, Satsangi, Doughty, Courtney (2013)</td>
<td>( N = 3 ) N(ASD) = 3 Ages = 6, 7, 10</td>
<td>Autism Clinic</td>
<td>Using concrete (physical objects) and virtual (3D objects on computer) manipulatives to help students acquire single-digit and double-digit subtraction skills</td>
<td>Virtual or concrete manipulative sessions consisting of ten trials with five subtraction problems each</td>
<td>Single Subject Alternating Treatment Design</td>
<td>Percentage of accurately completed subtraction problems and percentage of subtraction problem steps completed individually</td>
<td>Use of concrete and virtual manipulatives</td>
<td>All three participants demonstrated an increase in correctly completed subtraction problems using both the concrete and virtual manipulatives (the virtual manipulatives proved slightly more effective), as well as an increase in independent completion of problems</td>
</tr>
<tr>
<td>Burton, Anderson, Prater, and Dyches (2013)</td>
<td>( N = 4 ) N(ASD) = 3 Ages = 13, 13, 14, 15</td>
<td>Junior High School</td>
<td>Using a video self-modeling technique to teach money estimation of a given item and estimation of the amount to receive in change</td>
<td>Two daily sessions of intervention testing to complete five math problems with the iPad, occurring four days each week</td>
<td>Multiple-Baseline Across-Participants Design</td>
<td>Percentage of accurately completed money computational problems</td>
<td>Use of video self-modeling on the iPad</td>
<td>All four of the participants demonstrated an improvement in math skill performance after implementation of the video self-monitoring</td>
</tr>
<tr>
<td>Cihak and Foust (2008)</td>
<td>( N = 3 ) N(ASD) = 3 Ages = 7, 7, 8</td>
<td>Resource room at Elementary School</td>
<td>Using number line and touch-point strategies to solve single-digit addition mathematics problems</td>
<td>Two daily 5–20 minute sessions to complete a worksheet of ten single-digit mathematics problems using either touch points or a number line</td>
<td>Alternating Treatments Design</td>
<td>Percentage of correct single-digit addition math problems completed</td>
<td>Use of a touch points and use of a number line</td>
<td>All three students demonstrated that touch-point strategy was more successful in teaching single-digit addition skills than the number line strategy</td>
</tr>
<tr>
<td>Fletcher, Boon, and Chihak (2010)</td>
<td>( N = 3 ) N(ASD) = 2 Ages = 13, 13, 14</td>
<td>Self-contained classroom at Middle School</td>
<td>Teaching single-digit mathematics problems using TOUCHMATH, a multisensory mathematics program, and a number line</td>
<td>Two daily 5–15 minute sessions to complete a worksheet of ten single digit mathematics problems using either the “touch points” or number line strategy</td>
<td>Alternating Treatments Design</td>
<td>Percentage of single-digit mathematics problems answered correctly</td>
<td>Use of the TOUCHMATH program using “touch points” and the number line strategy</td>
<td>TOUCHMATH strategy was more effective and efficient in teaching single-digit addition problems compared to the number line strategy</td>
</tr>
<tr>
<td>Rockwell, Griffin, and Chihak (2011)</td>
<td>( N = 1 ) N(ASD) = 1 Age = 10</td>
<td>Author’s Home Office</td>
<td>Using schematic diagrams to solve group, change, and compare addition or subtraction word problems</td>
<td>Problem solving probe sessions containing two group, two change, and two compare word problems</td>
<td>Multiple Probes Across Behaviors Single-Case Design</td>
<td>Percentage of word problems completed accurately</td>
<td>Use of schematic diagrams to solve group, change, and compare word problems</td>
<td>The participant successfully increased her ability to complete single-step addition and subtraction word problems</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Setting</td>
<td>Intervention</td>
<td>Intervention Testing</td>
<td>Design</td>
<td>Dependent Variables</td>
<td>Independent Variables</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------</td>
<td>--------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Waters and Boon (2011)</td>
<td>Self-contained special education mathematics classroom at Public High School</td>
<td>Teaching 3-digit money computational subtraction problems with regrouping using the TouchMath program</td>
<td>Ten 3-digit money computational problems solved by regrouping using the touch-point strategy</td>
<td>Multiple-Probe Across Participants Design</td>
<td>Percentage of correct 3-digit money computations performed</td>
<td>Use of TouchMath program using &quot;touch points&quot; and regrouping</td>
<td>All three participants increased acquired skills to subtract 3-digit mathematics operations using money computations</td>
<td></td>
</tr>
<tr>
<td>Benda and Kubina Jr. (2010)</td>
<td>Resource room at Middle School</td>
<td>Using high preference math tasks to increase academic compliance and completion of low-preference math tasks</td>
<td>Ten cards containing two high-preference problems to be completed prior to one low-preference problem</td>
<td>ABAB Design</td>
<td>Lack of initiation of a three-digit by three-digit missing addend problem</td>
<td>Using a high preference intervention to complete ten test cards containing two three-digit by three-digit addition problems followed by one missing addend problem</td>
<td>The student took less time to begin low-preference math problems that were stated after high-preference math problems</td>
<td></td>
</tr>
<tr>
<td>Cihak and Grim (2008)</td>
<td>Resource room at High School, school bookstore, and local department store</td>
<td>Using counting-on math technique with the next-dollar strategy to increase independent purchasing skills</td>
<td>1) Classroom Phase: Two daily sessions containing ten trials of ten problems to be completed using the counting-on strategy 2) Bookstore and Community Phases: One daily session containing three purchasing trials</td>
<td>Multiple-Probe Design</td>
<td>Percentage of independent purchases completed accurately</td>
<td>Use of counting-on and next-dollar math strategies to enhance purchasing skills</td>
<td>All four students successfully learned the counting-on and next-dollar techniques and were able to apply these skills to community settings</td>
<td></td>
</tr>
<tr>
<td>Hua, Morgan, Kaldenburg, and Goo (2012)</td>
<td>University classroom</td>
<td>Using a three-step cognitive strategy (TIP) for calculating tip and total bill for young adults with intellectual disabilities</td>
<td>Six instructional stages including a final stage of ten tip and total bill problems to be completed independently by the students</td>
<td>Pre- and Posttest NonEquivalent Groups Design</td>
<td>Total number of tip and total bill calculations completed accurately</td>
<td>Use of TIP method</td>
<td>The experimental group successfully increased their ability to calculate tip and total bill using the TIP strategy, showing much higher posttest result than the comparison group</td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Participants in Experiment (Total Number of Participants / Number of Participants with ASD, Ages)</td>
<td>Setting</td>
<td>Intervention</td>
<td>Intervention Testing</td>
<td>Design</td>
<td>Dependent Variables</td>
<td>Independent Variables</td>
<td>Results</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------</td>
<td>--------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Rapp, Marvin, Nystedt, Swanson, Paananen, and Tabatt (2012)</td>
<td>N = 4 N(ASD) = 2 N(MMR) = 1 Ages = 7, 8, 9, 12</td>
<td>Individual classroom at School or at Home</td>
<td>Using response repetition as an error-correction technique to increase students’ ability to complete math facts and math computation</td>
<td>Two to four weekly sessions ranging from 15–20 minutes each, containing 2–5 sets of flash cards or worksheets per session</td>
<td>Nonconcurrent and Concurrent Multiple Baseline Designs</td>
<td>Percentage of mathematic problems solved accurately</td>
<td>Use of Response Repetition technique</td>
<td>Three of four participants demonstrated improvement on targeted math problems, and two of two participants showed improvement on other forms of math problems</td>
</tr>
<tr>
<td>Whitby (2013)</td>
<td>N = 3 N(ASD) = 3 Ages = 7th grader, 8th grader</td>
<td>Individual Public Middle School classrooms</td>
<td>Teaching math word problems using the 7 cognitive strategies and 3 metacognitive strategies from the Solve It! Problem Solving Routine</td>
<td>Minimum of five training sessions containing 3–5 math word problems; five acquisition condition sessions containing five math word problems</td>
<td>Multiple Baseline Across Participants Design</td>
<td>Percentage of correct word problems completed</td>
<td>Use of Solve It! Problem Solving Routine curriculum</td>
<td>All three participants learned to use problem-solving skills to accurately complete math word problems, as demonstrated through the increased percentage of correct word problem responses</td>
</tr>
</tbody>
</table>

Note: Table 1 lists the reviewed articles alphabetically based on Visual Representations or Cognitive Strategy Intervention type

* Excludes comparison group participants
therefore may have useful application to students with ASD.

In all, 34 students comprised the participants in the studies included in the review, with 28 participants diagnosed with an ASD (82.35%). The participants with ASD included three students diagnosed with Asperger’s Syndrome (8.82%), one student diagnosed with Pervasive Development Disorder (2.94%), and 17 students diagnosed with Autism (50.00%). Participants ranged from 6 to 22 years, and included students attending elementary school through individuals attending post-secondary programs for young adults with disabilities. A summary of participant characteristics is provided in Table 2. Moreover, although no specific date restrictions were utilized in the search, studies located in the review were published between 2008 and 2013. Specifically, there were three studies from 2012, two from each of the remaining years, and no studies from 2009.

### TABLE 2

<table>
<thead>
<tr>
<th>Student Diagnosis</th>
<th>Number of Students with Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe LD/ADHD</td>
<td>1</td>
</tr>
<tr>
<td>ID</td>
<td>1</td>
</tr>
<tr>
<td>ASD and ID</td>
<td>2</td>
</tr>
<tr>
<td>ASD</td>
<td>26</td>
</tr>
<tr>
<td>-Autism</td>
<td>9</td>
</tr>
<tr>
<td>-Autism and MID</td>
<td>7</td>
</tr>
<tr>
<td>-Autism and PDD</td>
<td>1</td>
</tr>
<tr>
<td>-AS</td>
<td>1</td>
</tr>
<tr>
<td>-AS and ID</td>
<td>1</td>
</tr>
<tr>
<td>-AS and MID</td>
<td>1</td>
</tr>
<tr>
<td>-ASD</td>
<td>6</td>
</tr>
<tr>
<td>MID</td>
<td>2</td>
</tr>
<tr>
<td>MR</td>
<td>1</td>
</tr>
<tr>
<td>MR and ID</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total Number of Participants</strong></td>
<td><strong>34</strong></td>
</tr>
</tbody>
</table>

Table Abbreviations:
- LD = Learning Disabilities
- ADHD = Attention Deficit Hyperactivity Disorder
- ID = Intellectual Disabilities
- ASD = Autism Spectrum Disorder
- MID = Mild Intellectual Disabilities
- MR = Mental Retardation
- PDD = Pervasive Development Disorder
- AS = Asperger Syndrome

Interventions Using Visual Representations

The six studies incorporating visual representation approaches consisted of three interventions using touch point (Cihak & Foust, 2008; Fletcher et al., 2010; Waters & Boon, 2011), one intervention using video self-modeling on an iPad (Burton et al., 2013), one intervention using virtual and concrete manipulatives (Bouck et al., 2014), and one intervention using schematic diagrams to solve word problems (Rockwell et al., 2011).

Of the visual representation interventions reviewed, three studies used the “touch-point” technique to teach mathematic skills to students with ASD (Cihak & Foust, 2008; Fletcher et al., 2010; Waters & Boon, 2011). Two of these interventions investigated the use of touch-points in comparison to a number line when teaching single-digit mathematics problems. During the first of these two studies (Cihak & Foust, 2008), instruction was provided to three students with autism. Students were presented a number line labeled with numbers zero through twenty and taught how to use the tool by moving their fingers along the line in correspondence to the numbers in each problem. Touch point instruction included teaching students the dot positions for the numbers one through nine, instructing students to count out loud the dots of the two given numbers, and writing the final number said aloud. Daily intervention sessions consisted of students performing a worksheet with ten single-digit math problems, alternating between using the touch point and number line methods. Results showed that the touch point method was more successful than the number line, and participants averaged 72% of the problems solved correctly using touch point, while the number line strategy averaged 17% of correctly solved problems (Cihak & Foust, 2008).

During the second number line and touch point intervention (Fletcher et al., 2010), three students with moderate intellectual disabilities and two diagnosed with ASD, were taught single-digit mathematics problems. The students were taught the same procedures to utilize the number line and touch point methods as was used in the previous intervention. Each intervention session required participants to solve ten single-digit...
addition problems using either the number line or touch point method. Consistent with the prior study that implemented these two strategies, the touch point method proved to be more effective in teaching single-digit math problems. Results show that participants solved an average of 92% of the problems correctly using the touch points, whereas they averaged 30% of problems using the number line (Fletcher et al., 2010).

Waters and Boon (2011) provided instruction to three students with mild intellectual disabilities, with two having an additional diagnosis of ASD. This third visual representation study used the touch-point strategy to teach three-digit money computational subtraction problems with regrouping. After mastering the TouchMath program’s procedure for using touch points, students were taught the method for solving subtraction problems with regrouping. This method follows a similar procedure to the addition touch-point strategy, but requires students to also count backwards and borrow numbers when regrouping. Results suggest that the touch point strategy is effective as demonstrated by each student’s increase in ability to solve three-digit subtraction problems with regrouping, a skill that each student lacked prior to the intervention (Waters & Boon, 2011).

Burton et al. (2013) provided training to three students with autism and one student diagnosed with an intellectual disability. Researchers implemented video self-modeling to teach money estimation of a given item, as well as the estimated amount to receive in change for that item. Prior to the intervention, a video was recorded of each student, in which the participant used a script featuring the seven steps of task analysis to solve a math problem. This procedure was repeated five times, resulting in five recorded videos. During the intervention, the student watched the video to observe himself complete the problem, and was able to pause, fast-forward, or rewind the recording as he completed the same problem on paper. The observed functional relationship between video self-modeling and performance supports the idea that this study was successful in improving the money skills of students with ASD (Burton et al., 2013).

Rockwell et al. (2011) implemented an additional visual representation intervention using schematic diagrams to help a student with ASD solve group (featuring two smaller parts combined to form one larger group), change (featuring a beginning amount, change amount signified by an action, and an ending amount), and compare (featuring a larger amount, smaller amount, and a difference) addition or subtraction word problems. The increased percentage of correct word problem responses in this study suggested that this intervention was effective (Rockwell et al., 2011).

A final intervention that utilized virtual and concrete manipulatives compared the effectiveness of these two strategies when implemented by students with ASD. Bouck et al. (2014) implemented this intervention with three students with ASD. Intervention trials alternated between students using either concrete blocks or an online computer program to solve subtraction problems. Although students increased the number of successfully completed single- and double-digit math problems using both strategies, the virtual manipulatives technique proved slightly more effective (Bouck et al., 2014).

Overall, visual representations were effective strategies in teaching students with ASD mathematic skills. Authors reported gains in single-digit addition problems (Cihak & Foust, 2008; Fletcher et al., 2010), subtraction problems (Bouck et al., 2014; Waters & Boon, 2011), money estimation (Burton et al., 2013), and addition/subtraction word problems (Rockwell et al., 2011). However, a limitation to the studies may include the fact that the majority of the visual representation interventions (Bouck et al., 2014; Burton et al., 2013; Rockwell et al., 2011; Waters & Boon, 2011) featured participants who received mathematic instruction outside of the general education setting, likely preventing them from accessing the mathematic content taught according to general education standards.

Interventions Using Cognitive Strategy Instruction

The five studies implementing cognitive strategies included two interventions citing specific use of cognitive or meta-cognitive strategies (Hua et al., 2012; Whitby, 2013), one intervention using response-repetition as an error-correction procedure (Rapp et al.,
one intervention using counting-on and next-dollar strategies (Cihak & Grim, 2008), and one intervention using a high-preference strategy to increase completion of low-preference tasks (Banda & Kubina Jr., 2010).

Two of the reviewed interventions used cognitive strategies to aid students in memorizing the steps required to successfully complete assigned mathematics problems (Hua et al., 2012; Whitby, 2013). Whitby (2013), the first of these interventions, achieved the goal of teaching students to accurately complete mathematic word problems by instructing students to memorize and apply Solve It! Problem Solving Routine’s seven cognitive steps (i.e., read-paraphrase-visualize-hypothesize-estimate-compute-check). This study also required the three participants with ASD to utilize the program’s three meta-cognitive strategies, which included self-management, self-questioning, and self-evaluation. During the intervention, all three students solved more problems accurately than their peers, suggesting the Solve It! method to be an effective strategy for teaching problem solving skills (Whitby, 2013).

The second intervention (Hua et al., 2012) taught five students with either ASD or intellectual disability to utilize a three-step cognitive strategy, TIP, to calculate tip and total bill. Results of this cognitive strategy-based intervention showed the method was effective in helping students complete tip and total bill computations, as well as in aiding students to accurately apply word problem solving techniques. Additionally, these results confirm that cognitive strategies are beneficial in helping students with ASD memorize processes required to successfully solve mathematic problems (Hua et al., 2012).

Rapp et al. (2012) utilized a cognitive strategy intervention of response repetition as an error-correction technique to increase computation of math facts. The study included four students diagnosed with disabilities, with two students with ASD, one with Asperger syndrome, and one with a moderate intellectual disability. During the intervention phase, a trainer presented math problems and provided verbal praise for accurate answers. If the student provided an incorrect answer, the trainer prompted the student to verbally state the correct answer five times. Response repetition was demonstrated to be a successful cognitive strategy for three of four participants (Rapp et al., 2012).

Cihak and Grim (2008) studied four students with moderate intellectual disabilities and ASD. Students were taught counting-on and next-dollar math techniques to increase independent purchasing skills. With the next-dollar strategy, students were taught to pay with one more dollar than asked for by the sales associate. The counting-on method taught students to count on from the first addend to obtain the sum. Results of this cognitive strategy showed that students developed and maintained independent purchasing skills, and were able to apply these skills to community settings (Cihak & Grim, 2008).

The final cognitive strategy study (Banda & Kubina Jr., 2010) required a middle school student with ASD to complete high-preference mathematics problems followed by the completion of low-preference problems, with the goal of increasing the initiation of the low-preference problems. Each intervention phase featured a stack of ten cards containing two high-preference problems followed by one low-preference problem for the student to complete. The results of this intervention showed that the student not only took less time to begin low-preference math problems, but also increased performance on the low-preference three-digit by three-digit missing addend problems (Banda & Kubina Jr., 2010).

In summary, cognitive strategy interventions were successful in increasing students’ mathematics skills and understanding. Authors discussed improvements among students with ASD in solving word problems (Whitby, 2013), money computational problems (Cihak & Grim, 2008; Hua et al., 2012), three-digit by three-digit missing addend problems (Banda & Kubina Jr., 2010), and addition, subtraction, and multiplication problems (Rapp et al., 2012).

Setting of Intervention, Daily Instruction, and Mathematics Instruction

As part of this analysis, studies were reviewed to determine the setting of the interventions themselves, the daily instructional setting of
participating students, and the daily mathematics instructional setting for each participant.

**Location of intervention.** Every article in this review described studies that were conducted in secluded areas outside of the general education setting. These settings included school resource classrooms (Banda & Kubina Jr., 2010; Gihak & Foust, 2008; Gihak & Grim, 2008), self-contained classrooms (Burton et al., 2013; Hua et al., 2012; Waters & Boon, 2011; Fletcher et al., 2010), individual secluded classrooms (Rapp et al., 2012; Whitby, 2013), a student or author’s home (Rapp et al., 2012; Rockwell et al., 2011), and an autism clinic (Bouck et al., 2014).

**Location of majority of daily instruction.** The studies that did provide information regarding the setting of daily instruction indicated that 4 students received most of their daily instruction in the general education setting (Rockwell et al., 2011; Whitby, 2013), while 25 spent the majority of the day in a self-contained classroom, resource classroom, or other secluded setting (Bouck et al., 2014; Burton et al., 2013; Gihak & Foust, 2008; Gihak & Grim, 2008; Fletcher et al., 2010; Hua et al., 2012; Waters & Boon, 2011).

**Location of mathematics instruction.** The analysis also indicated that 3 students received mathematics instruction in the general education setting (Whitby, 2013), while 16 of the participating students received mathematics instruction in a self-contained classroom, resource classroom, or other secluded setting (Bouck et al., 2014; Burton et al., 2013; Hua et al., 2012; Rockwell et al., 2011; Waters & Boon, 2011). This data suggests that although an increasing number of students with ASD are being included in the general education setting, the majority of the current mathematics interventions are comprised of students who received mathematics instruction in a self-contained or other segregated setting, which may in part be a function of their overall intellectual levels. Nonetheless, these students’ access to the general education curriculum was relatively limited, preventing them from acquiring math content according to general education curricular standards.

**Mathematical Skill Targeted: Academic or Functional**

Reviewed articles were also analyzed based on the targeted skill type, either academic or functional. Academic interventions were identified as having the goal of strengthening students’ academic skills and understanding of mathematical concepts. Functional interventions were identified as having the goal of improving life skills. Academic interventions were included in seven of the reviewed articles (Banda & Kubina Jr., 2010; Bouck et al., 2014; Gihak & Foust, 2008; Fletcher et al., 2010; Rapp et al., 2012; Rockwell et al., 2011; Whitby, 2013), while functional interventions categorized four of the articles (Burton et al., 2013; Gihak & Grim, 2008; Hua et al., 2012; Waters & Boon, 2011).

A closer inspection of the seven academic interventions revealed only two interventions focused on improving mathematical word problem solving skills of students with ASD (Rockwell et al., 2011; Whitby, 2013). Three studies were implemented with the goal of teaching addition facts, specifically single-digit addition problems (Cihak & Foust, 2008; Fletcher et al., 2010; Hua et al., 2012; Waters & Boon, 2011). Additionally, one study focused on teaching single-digit and double-digit subtractions skills (Bouck et al., 2014), while the final study taught addition skills to two participants, addition and subtraction skills to another participant, and multiplication skills to the fourth participant (Rapp et al., 2012).

The four studies categorized as functional interventions focused on improving the money skills of students with ASD. The first of these studies taught three digit money computational subtraction problems using regrouping and touch point (Waters & Boon, 2011). The second intervention used the counting-on math technique with the next-dollar strategy to increase independent purchasing skills (Cihak & Grim, 2008). A three-step cognitive strategy (TIP) for calculating tip and total bill was taught during the third intervention (Hua et al., 2012). The final intervention used video self-modeling to teach money estimation of a given item, and the estimated amount to receive in change for
that item (Burton et al., 2013). These four interventions were all effective in teaching basic mathematical processes to improve the money skills of students with ASD. Moreover, results showed that many participants were able to apply the acquired techniques in community settings.

For the functionally driven interventions, although students’ successful implementation of critical life skills in settings beyond the classroom supports their ability to acquire, maintain, and generalize mathematic material, mathematical processes targeted were generally low level. Additionally, despite the fact that more of the reviewed studies focused on academic rather than functional skills, the academic interventions themselves also mainly targeted low-level mathematical content, including basic addition and subtraction skills. Without understanding of common advanced mathematic knowledge and processes (e.g., multiplication, division, or algebraic skills) students with ASD, particularly those less affected by intellectual disability, may not be prepared for the more rigorous academic content of the general education classroom.

**Discussion and Implications**

Given their increasing prevalence rates and more frequent placement in general education classrooms, there exists a growing need to teach general education curricular academics to students with ASD, particularly in the area of mathematics. Correspondingly, targeted attention to mathematics related IEP goals and related interventions is essential to meeting individual needs as many students with ASD require support to develop the conceptual and procedural understanding necessary to engage in mathematical computation and problem solving successfully. From this review, it is clear that research evaluating mathematics interventions for learners with ASD is preliminary and may therefore be insufficient to effectively guide teacher practice. The results of this review point to the need for future research to extend the literature base reviewed here. Future research should include investigations specific to students with ASD that are more specifically focused on the higher-order skills consistent with NCTM recommendations and that maximize students’ access to the general education curriculum as is appropriate. Moreover, evidence-based interventions found to be effective with learners with other types of disabilities (e.g., learning disability) should be systematically tested for possible usefulness for students with ASD, in accord with the quality indicators for research methodology and evidence-based practices established by the Council for Exceptional Children’s Division for Research (Odom et al., 2005).

Future single-subject design investigations should meet standard expectations for rigor by including the actual measures used in the mathematics activities, such as pre and post-test measures and generalization measures that establish a pattern of academic behavior or instances of deviation from a pattern, thereby producing more detailed information for evaluation by researchers and educators alike. Moreover, reporting how intervention fidelity is assessed is also critical. Encouragingly, among the studies reviewed herein, ten provided specific information on the extent to which interventions were implemented with fidelity (Banda & Kubina Jr., 2010; Bouck et al., 2014; Burton et al., 2013; Cihak & Foust, 2008; Cihak & Grim, 2008; Fletcher et al., 2010; Hua et al., 2012; Rockwell et al., 2011; Waters & Boon, 2011; Whitby, 2013), with nine of these studies recording a treatment integrity of at least 95% (Banda & Kubina Jr., 2010; Bouck et al., 2014; Burton et al., 2013; Cihak & Foust, 2008; Cihak & Grim, 2008; Fletcher et al., 2010; Hua et al., 2012; Rockwell et al., 2011; Whitby, 2013). However, multiple articles described treatment fidelity being calculated for only a specified number of intervention sessions, instead of being recorded for all implemented sessions (Cihak & Grim, 2008; Fletcher et al., 2010; Waters & Boon, 2011). This lack of consistent tracking of treatment fidelity is a possible limitation of the studies reviewed.

Of the eleven studies reviewed, six interventions demonstrated the skill being taught to mastery in at least three consecutive intervention or post-intervention sessions (Bouck et al., 2014; Cihak & Foust, 2008; Cihak & Grim, 2008; Fletcher et al., 2010; Rockwell et al., 2011; Waters & Boon, 2011). This mastery criterion was often recorded as 100% accuracy.
in problems completed during a session. In the remaining five studies, data showed that students improved their mathematics skills because of the intervention, but all participants did not demonstrate mastery in three consecutive sessions as is often standard in research practice (Banda & Kubina Jr., 2010; Burton et al., 2013; Hua et al., 2012; Rapp et al., 2012; Whitby, 2013). Although these latter studies recorded significant improvements as a result of the interventions, they could be strengthened by requiring students to achieve at least three consecutive sessions of mastered criterion (e.g. 100% performance on the intervention testing problems each session), which results in a greater possibility of skills being maintained over time. Additionally, nine of the reviewed articles included a post intervention assessment or follow-up to evaluate maintenance of the intervention skills and to determine the prolonged treatment effect (Bouck et al., 2014; Burton et al., 2013; Cihak & Foust, 2008; Cihak & Grim, 2008; Fletcher et al., 2010; Hua et al., 2012; Rockwell et al., 2011; Waters & Boon, 2011; Whitby, 2013). Future research should continue to focus on meeting these important standards for rigor in their design, implementation, and analysis (Horner et al., 2005).

Mathematics instruction in special education has been typified to a considerable extent by its focus on rote memorization of facts and computational skills, rather than on developing and expanding mathematical skills to be applied in authentic problem situations (Woodward & Montague, 2002). On the other hand, the NCTM has emphasized meaningful instruction for all students in order to assist them in developing understanding of mathematical content and procedures to engage in problem solving (NCTM, 2002). Few studies examined in this review referenced NCTM standards, and the majority targeted relatively basic mathematical computation skills. Because NCTM standards are for all students, future interventions for students with ASD should target such mathematics requirements. By developing interventions based on NCTM standards, students with ASD will gain stronger mathematics skills to enable them to be more successful in the general education setting. Furthermore, given that recent literature suggests nearly 40% of individuals with ASD have average to above average levels of intelligence and that ASD is becoming less strongly associated with intellectual disability than traditionally held (Croen, Grether, Hoogstrate, & Selvin, 2002), research interventions should reflect a focus on higher-order skills so that teachers can be prepared to provide access to high-quality mathematics instruction for students with ASD included in their classrooms.

Math interventions with a strong evidence base should be applied in settings serving students with ASD. Recent reviews of mathematics interventions indicate that there is empirical support for cognitive strategies including verbal strategy instruction (Montague, 2008) and visual-spatial strategies (van Garderen, 2007) as well as meta-cognitive strategies (Xin & Jitendra, 1999) for enhancing students’ mathematic performance. Such studies have been limited primarily to students with diagnosed learning disability and could be expanded to evaluate their effectiveness for students with ASD, many of who also contend with mathematics learning difficulties. Based on this review, there is emerging evidence to support cognitive and visually based strategies for learners on the autism spectrum but additional studies are needed.

**Limitations**

One specific limitation associated with this literature review is publication bias, which is a possible confound in most literature syntheses (Reichow & Volkmar, 2009). That is, there may be systematic differences between studies that were selected for inclusion and those that were excluded from the review. Although an extensive literature search was conducted utilizing multiple methods and sources, the narrow inclusionary criteria that required studies to be peer-reviewed may have created study selection bias. In addition, the recency of publication for the studies included in this review (i.e., 2008–2013) may also have narrowed its scope, though given the relatively current rise in identification of students with ASD, it is not surprising that most intervention studies are also correspondingly recent.
Conclusion

Poor post secondary outcomes for youth with ASD point to the need for a re-evaluation of the quality and quantity of academic preparation individuals with ASD receive throughout their schooling (Fleury et al., 2014). Although a substantial number of students with ASD contend with mathematical difficulties, research addressing the mathematical knowledge and skills of children and youth with ASD is in its infancy and is generally limited. Researchers should continue investigating this line of inquiry with a focus on interventions that stress conceptual knowledge and problem solving and that attend to the principles, standards, and skill areas described by the NCTM. In the interim, teachers should provide mathematics instruction from the earliest grades that is consistent with recommendations made by the NCTM and the existing research on effective mathematics interventions for students with ASD, while educational researchers expand their investigations to apply other evidence-based mathematics interventions in settings serving such students. Based on this review, emerging evidence suggests that learners with ASD can make gains in mathematics skills when targeted interventions are implemented.

References


Received: 30 January 2014
Initial Acceptance: 1 April 2014
Final Acceptance: 1 May 2014
Is There Really a Difference? Distinguishing Mild Intellectual Disability from Similar Disability Categories

Emily C. Bouck  
Michigan State University

Rajiv Satsangi  
Purdue University

Abstract: Students with mild intellectual disability generally garner less individual attention in research, as they are often aggregated with students with moderate and severe intellectual disability or students with other high incidence disabilities. This study used the National Longitudinal Transition Study-2 (NLTS2) to look at the personal characteristics and educational curriculum offered to students with mild intellectual disability during high school, and compared their data to other disability groups to determine if students with mild intellectual disability are unique. Results from the study suggested students with mild intellectual disability differ in cognitive skills and aptitude, and are generally afforded greater classroom accommodations than high incidence disability categories. Statistically significant differences found on direct assessment data and self-determination measures support the notion students with mild intellectual disability are qualitatively and quantitatively unique, and deserve unique consideration from other disability categories.

Over the past few decades, the category of mild intellectual disability gradually transformed in terms of prominence and attention within special education research and practice (Fujiura, 2003; Polloway, 2006). Although once considered the largest disability category within special education, the category of mild intellectual disability experienced a decline in relevance in recent years (Bouck, 2007, 2012; Edgar, 1987; Gargiulo, 2012; Hourcade, 2003). Although the decline may be due to many reasons, often cited explanations include the lack of clarity on the part of educators on how to identify the disability as well as reluctance by educators to actually label a child as having a mild intellectual disability (Beirne-Smith, Patton, & Kim, 2006; Harris, 2006). The number of students identified as having a mild intellectual disability fluctuates greatly across states, with the category accounting for large portions of the student population in some states, and virtually no students in other states (Polloway, Lubin, Smith, & Patton, 2010).

The disability category of mild intellectual disability is categorized as below-average intellectual functioning, in conjunction with related struggles in multiple adaptive skill areas, including communication skills, social skills, self-care, functional academics, home or school living, community use, self direction, health and safety, work, and leisure skills (Polloway, Patton, Smith, & Buck, 1997). Individuals with mild intellectual disability are generally characterized as having IQ scores ranging from 55–70 (Bouck, 2012; Schalock et al., 2010). Characteristics for students with mild intellectual disability include limited attention span (Dunn, 1973; Kirk, 1972; Thomas, 1996); and difficulty with transferring, processing, generalizing, and recalling information (Belmont, 1966; Spitz, 1973, Stephens, 1972). Academically, students with mild intellectual disability typically experience delays in developing foundational language, reading, and mathematics skills (Taylor, Richards, & Brady, 2005).

Although the current prevalence rate of students with mild intellectual disability is not known, under IDEA 443,054 students between the ages of 3 through 21 were identified as having an intellectual disability in 2011—the latest data available (TA&D, n.d.). However, King, Toth, Hodapp, and Dykens (2009) suggested about 85% of individuals identified...
with intellectual disability can be considered as having mild intellectual disability; Drew and Hardman (2007) suggested 90%. In education, the category of mild intellectual disability is often considered a high-incidence disability, meaning that its incidence rate occurs more often in the population (Gage, Liermeimer, & Goran, 2012). The disability categories considered high incidence disabilities (e.g., students with learning disabilities, students with emotional/behavior disorders, students with ADHD, and students with mild intellectual disability) are sometimes referred to as mild disabilities, despite students within those categories (e.g., students with mild intellectual disability) not always presenting mild needs (Polloway, 2004).

As previously noted, one reason for the decline in identification and subsequent attention to students with mild intellectual disability includes the aggregation of students with mild intellectual disability with other disability categories, such as moderate intellectual disability or learning disability. Despite historical differences in IQ scores and adaptive behavior needs, as well as educational programming, students with mild intellectual disability are losing their identity within the larger category of students with intellectual disability (Bouck, 2004; Hardman, Drew, & Egan, 2002). With distinct characteristics of each subcategory of intellectual disability, educational practices and implications of research would presumably differ across each. However, over the past three decades, research on mild intellectual disability significantly decreased, with limited literature found explicitly targeting this group’s educational programming in the areas of curriculum and training (Bouck, 2004, 2007; Edgar, 1987; Keogh & MacMillan, 1996). In its place, research in the field of intellectual disability appears to target students with moderate and severe intellectual disability or aggregate the categories of intellectual disability when reporting findings (Bouck, 2004). Such trends in research do not consider the unique learning characteristics and outcomes of each subcategory of intellectual disability–instead opting to treat mild, moderate, and severe intellectual disability as part of a broader classification of disability (Bouck, 2004; Edgar, 1987; Keogh & MacMillan, 1996).

Similar to the trend of aggregating the categories of intellectual disability, the decline in the identification of students with mild intellectual disability may also be attributed (a) to the obstacles encountered when discriminating amongst categories of students presenting characteristics associated with high incidence disabilities (e.g., learning disability, emotional/behavioral disability, ADHD) as well as students functioning on the margins of what may be considered a disability group (e.g., low academic achievers or borderline students; Gresham, MacMillan, & Bocian 1996; MacMillan, Gresham, Bocian, & Lambros, 1998), or (b) a preference to identify students with mild intellectual disability as another high incidence disability (Beirne-Smith et al., 2006). Thus, in research, students with mild intellectual disability are often aggregated with other high incidence disability categories, and results are subsequently reported for the aggregate group (Polloway, 2004; 2005). Evidence suggests this aggregation is also occurring in practice (Bouck, 2007; Polloway, 2004; 2006), with a general preference not to label a child as having a mild intellectual disability (Beirne-Smith et al., 2006).

Previous researchers examined students diagnosed with high incidence disabilities to determine similarities and differences (Gage et al., 2012; Sabornie, Evans, & Cullinan, 2006); the literature suggested conflicting results. Over thirty years ago, researchers speculated the behavioral and academic characteristics of high incidence disabilities were more similar than different (Hallahan & Kauffman, 1976, 1977). Gresham et al. (1996) compared children identified with mild intellectual disability, learning disabilities, and low achievement across a range of cognitive, achievement, social-behavioral, and school history variables and found significant differences between the three groups on measures of aptitude and achievement. Among their findings, the authors showed students with mild intellectual disability performed poorest across all academic areas when compared to students with learning disabilities and low achievement. Yet, their data also suggested virtually no variation between the three groups across variables such as social skills, academically engaged time, emotional behavior, or school history. Despite average levels of discrepancy between...
the mild intellectual disability/learning disabilities group and mild intellectual disability/low achievement group, the authors concluded there were more similarities than differences overall between the three groups—with those differences not educationally relevant in terms of interventions, accommodations, or placement. The findings of Gresham et al. (1996) align with the suggestions offered nearly four decades ago by some in the field who championed a cross-categorical approach to teaching students diagnosed with high incidence disabilities (Hallahan & Kauffman, 1976, 1977). Recently, analyses of one school district in one state by Gage et al. (2012) confirmed the conclusion by Hallahan and Kaufman (1977) that non-categorical service delivery is appropriate.

In opposition to the idea that the characteristics of students with high incidence disabilities are predominantly homogeneous in nature, Sabornie et al. (2006) argued students with mild intellectual disability were substantially different with respect to lower IQ scores and lower academic achievement from students identified with learning disabilities and emotional-behavioral disabilities. Through a descriptive review of 34 studies comparing characteristics of high incidence disabilities, Sabornie et al. (2006) found significant differences between the three disability categories, and argued students with mild intellectual disability possessed considerably different cognitive and behavioral profiles, which must be taken into consideration when formulating academic programming.

The aggregation of students with mild intellectual disability—in research and practice today—is occurring without careful consideration as to whether the characteristics of students with mild intellectual disability differ qualitatively and quantitatively from students with moderate and severe intellectual disability or students with other high incidence disabilities, such as students with learning disabilities (Gresham et al., 1996; Polloway et al., 2010). As the category of mild intellectual disability loses prevalence in attention, the educational needs (i.e., curriculum, accommodations, post-school transitioning) of students with this disability may also be lost (Bouck & Joshi, 2012; Polloway, 2004, 2005). This study sought to investigate whether mild intellectual disability was in fact a qualitatively and quantitatively unique category. To address this issue, the authors identified the following research questions: (a) based on direct assessment data of academic content knowledge and self-determination measures, are students with mild intellectual disability significantly different from other disability groups such as moderate/severe intellectual disability, learning disability, or other high incidence disabilities (i.e., attention deficit hyperactivity disorder, emotional and behavioral disorder, speech and language disorder) ?, and (b) how do the educational curriculum, services, and accommodations offered to secondary students differ across the disability categories of mild and moderate/severe intellectual disability, learning disability, and other high incidence disabilities?

**Method**

This study used the National Longitudinal Transition Study-2 (NLTS2) database to conduct a secondary analysis of the cognitive and functional skills, classroom supports, and educational experiences of secondary high school students with mild intellectual disability in relation to students with moderate/severe intellectual disability, students with learning disabilities, and students with high incidence disabilities aggregated. The following section summarizes the NLTS2 and the participants and procedures used to conduct analysis. Additional information regarding this database can be found through their website (http://www.nlts2.org) and existing published work (Wagner, Kutash, Duchnowski, & Epstein, 2005).

**National Longitudinal Transition Study-2**

Created in 1985, The National Longitudinal Transition Study (NLTS) was a multiyear project funded by the Office of Special Education Programs and conducted by SRI International. With a focus on secondary students with disabilities receiving special education services, this project captured the characteristics of, and experiences encountered in secondary education, transition to post-school, and post-school outcomes for students with disabilities (Wagner, Newman, Cameto, Garza,
The National Longitudinal Transition Study-2 (NLTS2) was the follow-up to the NLTS. As a 10 year, government-sponsored project, the NLTS2 examined issues regarding secondary education, the transition to post-school, and post-school outcomes of students with disabilities (Newman, 2005). Beginning in the 2000–2001 academic school year, and ending in the 2008–2009 academic year, this project collected data through a variety of means, including: (a) parent and/or youth telephone interviews, (b) direct assessments of students, (c) teacher survey, (d) school program survey, (e) school information survey, and (f) student transcripts (SRI International, 2000b).

A two stage sampling process was used for the NLTS2 (SRI International, 2000a; Wagner et al., 2005). In total, 19,899,621 special education students from 12,435 state supported schools and Local Educational Agencies were selected to partake in the NLTS2 study (SRI International, n.d.). Factors such as geographic region, student enrollment within disability categories, and the socioeconomic status of each area were taken into account during the selection process. Students randomly chosen to participate were between the ages of 13 and 16, enrolled in seventh grade or higher, and received special education services (SRI International, n.d.). Selection was weighted towards older students (i.e., students who were 16 years of age), and done in a manner to ensure a standard error of 3.6% in the disability categories with the highest rates of diagnoses (i.e., learning disabilities, emotional/behavioral disorders, intellectual disability, speech and language impairments, other health impairments, and hearing impairments) (SRI International n.d.; Wagner et al., 2005).

Participants

To be included in this secondary analysis, students in the NLTS2 database met the following criteria: (a) identified as having a mild intellectual disability, moderate/severe intellectual disability, learning disability, or other high incidence disability (i.e., attention deficit hyperactivity disorder, emotional/behavioral disorder, speech and language disorder) in their IEP as reported in the school program database; (b) enrolled in school in wave 2 of data collection; and (c) received special education services while enrolled in school. Although analyses were conducted with students who met these criteria in the sample, all reported data were weighted using weights provided in the NLTS2 database to represent the number of students in the population (refer to Javitz & Wagner, 2003; Wagner et al., 2005 for more information on weighting the data). Data with low un-weighted counts were not reported in this analysis.

In total, 1,025,928 students were used for this secondary analysis (refer to Table 1 for participants’ demographic data, broken down by disability category). Of all participants, roughly two-thirds were male, with the majority identifying their ethnicity as Caucasian. Although distributed, the most frequent grade levels represented by participating students were 12th grade, followed second by 11th grade. Finally, the majority of students came from families whose reported annual family income was less than $25,000.

Data Collection

For this secondary analysis, researchers extracted data from the second of five waves of data collection of the NLTS2. Researchers used the Parent/Youth survey, the Students’ School Program survey, and direct assessment data collected for Wave 2 (i.e., the academic years of 2002–2003 and 2003–2004). The Parent/Youth survey in Wave 2 consisted of students completing a 60-minute phone interview. Interviews were conducted with a student’s parent if the student was unable to complete it, and a mail survey was conducted with either group if a phone interview was not feasible. Questions in the Parent/Youth survey focused on students’ lives outside of school, including household characteristics, non-school factors (e.g., extracurricular and community activities), family involvement, employment factors (e.g., type of work, salary), academic experiences, personal and social issues, citizenship, personal health issues, satisfactions, and behaviors (SRI International, 2000b).

A select teacher familiar with the student’s school program completed the mailed Students’ School Program survey. Surveys fo-
cused questions on the school program, special education services/accommodations, classroom supports, transition services, state and district assessments, student performance, and parent participation (SRI International, 2000b). Finally, researchers examined two types of direct assessment data. First, direct assessment data were collected individually for students on the following academic competencies: passage completion, applied problems, calculation, social science, science, and synonyms/antonyms standard scores using the Woodcock-Johnson III Tests of Achievement (SRI International, 2000b; Woodcock, Mather, & McGrew, 2001). The second form of direct assessment data examined was in reference to self-determination, relative to empowerment (Wehmeyer & Kelchner, 1995). Frequencies were measured on the following six measures: student’s ability to offer their opinions, make their own decisions, work hard to attain what they desire, make good choices, have their choices honored by others, and the ease with which they give up.

Procedure

To begin, researchers identified the relevant variables in each database and reduced the three databases to reflect only those chosen for analysis (refer to Table 2 for list of variables). Once the desired variables were chosen from the Parent/Youth, Direct Assessment, and the School Program databases, the unnecessary remaining variables were deleted. The variables from the Parent/Youth and Direct Assessment databases were then merged into the School Program database, which served as the main database. Once the databases merged, students whose primary disability did not fall into one of the following five categories were eliminated: mild intellectual disability, moderate/severe intellectual disability, learning disabilities, attention deficit

| TABLE 1 |
| Participant Demographic Data |

<table>
<thead>
<tr>
<th>Demographics</th>
<th>All (Overall)</th>
<th>MID (SE)</th>
<th>M/S ID (SE)</th>
<th>LD (SE)</th>
<th>Other High Incidence (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>66.9% (2.7)</td>
<td>58.9% (4.9)</td>
<td>66.9% (9.1)</td>
<td>67.7% (3.4)</td>
<td>68.2% (3.9)</td>
</tr>
<tr>
<td>Female</td>
<td>33.1% (2.7)</td>
<td>41.1% (4.9)</td>
<td>33.1% (9.1)</td>
<td>32.3% (3.4)</td>
<td>31.8% (3.0)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>25.0% (1.9)</td>
<td>17.4% (4.0)</td>
<td>12.5% (4.3)</td>
<td>27.2% (2.5)</td>
<td>27.4% (2.4)</td>
</tr>
<tr>
<td>17</td>
<td>31.5% (2.6)</td>
<td>22.7% (3.4)</td>
<td>15.8% (4.9)</td>
<td>36.7% (3.3)</td>
<td>34.4% (3.1)</td>
</tr>
<tr>
<td>18</td>
<td>22.7% (2.0)</td>
<td>26.3% (4.0)</td>
<td>17.9% (5.3)</td>
<td>23.3% (2.7)</td>
<td>22.6% (2.5)</td>
</tr>
<tr>
<td>19–20</td>
<td>20.7% (3.4)</td>
<td>33.5% (6.4)</td>
<td>53.7% (12.0)</td>
<td>12.8% (3.6)</td>
<td>15.6% (3.7)</td>
</tr>
<tr>
<td>Grade 9th</td>
<td>1.6% (0.5)</td>
<td>1.2% (0.6)</td>
<td>1.4% (0.8)</td>
<td>1.6% (0.6)</td>
<td>1.7% (0.6)</td>
</tr>
<tr>
<td>Grade 10th</td>
<td>18.1% (1.9)</td>
<td>18.9% (4.0)</td>
<td>14.2% (4.5)</td>
<td>18.3% (2.8)</td>
<td>18.3% (2.4)</td>
</tr>
<tr>
<td>Grade 11th</td>
<td>34.3% (2.3)</td>
<td>22.2% (3.4)</td>
<td>15.2% (4.6)</td>
<td>39.3% (3.3)</td>
<td>38.0% (2.8)</td>
</tr>
<tr>
<td>Grade 12th</td>
<td>40.6% (3.2)</td>
<td>46.5% (5.3)</td>
<td>27.2% (7.8)</td>
<td>40.3% (3.8)</td>
<td>41.0% (3.6)</td>
</tr>
<tr>
<td>Grade 13th</td>
<td>5.2% (1.8)</td>
<td>11.1% (2.8)</td>
<td>42.0% (14.7)</td>
<td>0.1% (0.1)</td>
<td>0.7% (0.4)</td>
</tr>
<tr>
<td>Income ≤ $25,000</td>
<td>41.9% (3.3)</td>
<td>41.3% (3.4)</td>
<td>29.2% (9.0)</td>
<td>44.8% (3.9)</td>
<td>43.3% (3.6)</td>
</tr>
<tr>
<td>$25,001–$50,000</td>
<td>26.9% (2.1)</td>
<td>31.8% (6.5)</td>
<td>22.1% (7.0)</td>
<td>27.6% (2.7)</td>
<td>26.7% (2.5)</td>
</tr>
<tr>
<td>&gt; $50,000</td>
<td>31.1% (3.4)</td>
<td>27.0% (5.0)</td>
<td>48.7% (14.2)</td>
<td>27.6% (3.5)</td>
<td>30.1% (3.5)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>62.6% (4.5)</td>
<td>65.0% (4.5)</td>
<td>45.2% (11.8)</td>
<td>60.5% (5.3)</td>
<td>63.9% (5.1)</td>
</tr>
<tr>
<td>African American</td>
<td>19.9% (2.8)</td>
<td>26.0% (4.1)</td>
<td>41.4% (14.7)</td>
<td>17.0% (2.7)</td>
<td>16.8% (2.9)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>15.1% (2.5)</td>
<td>7.1% (2.3)</td>
<td>11.3% (3.7)</td>
<td>19.5% (3.3)</td>
<td>16.7% (2.8)</td>
</tr>
</tbody>
</table>

Note: 13th grade represented students who received educational services beyond the 12th grade by their school.
hyperactivity disorder, emotional and behavioral disorder, and speech and language disorder. Attention deficit hyperactivity disorder, emotional and behavioral disorder, speech and language disorder were combined into a separate category labeled other high incidence disabilities. Researchers explicitly focused on students with learning disabilities for additional comparative analysis due to high prevalence rates of this disability category and often-cited commonalities to mild intellectual disability (Sabornie et al., 2006).

**Data Analysis**

This secondary analysis focused on descriptive statistics to examine variables most relevant to the research questions posed in this study. Researchers calculated frequency distributions for student demographic variables, including age, gender, ethnicity, disability, income, and grade level. The mean and range were also calculated for the students' age variable. Frequency distributions and standard error were also calculated for students' education.
tional programming, such as curriculum received and instructional environments, as well as the accommodations provided (i.e., extra time on testing, tests read aloud, modified tests, alternative tests, modified grading, slower paced instruction, and extra time on assessments).

To assess educational programming, data were collected on the amount of instruction students received in the following subject areas: life skills, language arts, social studies, mathematics, and science. Each of these five instructional areas was reported as a percent out of 100. The classroom environments in which students received instruction in each of these five subject areas were also measured. Students received curriculum instruction within a general education classroom, special education classroom, community setting, or through individual instruction. However, due to the small overall raw numbers of students who received instruction in a community setting or through individual instruction, such data were not reported in this study. An assessment of students’ functional skills was also gathered via a variable labeled “mental skills” in the Parent/Youth survey. The mental skills variable represented the sum of respondents’ evaluation of the student’s independent skills to tell time on a clock, read signs, count change, and use a phone. For the purposes of the secondary analysis, the average scores were reported. Each skill was assessed on a four point rating scale, ranging from, “not at all well” (one) to “very well” (four), resulting in an overall cumulative score from 4–16.

Finally, direct assessment data were analyzed. For the purposes of the secondary analysis, the average raw standard scores from the Woodcock Johnson Tests of Achievement (Woodcock et al., 2001) in the areas of passage completion, synonyms/antonyms, applied problems, calculation, social science, and science were reported; scores were centered on a mean of 100 and could range from 0–200. An Analysis of Variance (ANOVA) and Bonferroni post-hoc analyses were conducted to assess the statistical significance of mild intellectual disability against the other disability categories for each of the six direct assessment measures. Related, direct assessment data relative to self-determination (i.e., empowerment) were also examined. Assessed on a two-point rating scale, frequencies were measured through a student’s ability to offer their opinions, make their own decisions, work hard to attain what they desire, make good choices, have their choices be honored by others, and the ease with which they give up (Wehmeyer & Kelchner, 1995). For each question, participants selected a 1 or 2—with each option representing opposing sentiments. For example, when asked about decision-making, participants could select option 1, “I make my own decisions,” or option 2, “Other people make my decisions for me.” Statistical significance was determined using an F test—provided with the NLTS2 databases—for each of the six measures, comparing mild intellectual disability and the other three categories (i.e., moderate/severe intellectual disability, learning disabilities, and other high incidence disabilities; Wagner, Newman, Cameto, Levine, & Marder, 2007).

Results

Direct Assessment Data

Students with mild intellectual disability earned the following scores (reported as average raw standard scores) on six academic measures: passage completion ($\mu = 58.9; SE = 2.0$), applied problems ($\mu = 61.7; SE = 2.3$), calculation ($\mu = 65.2; SE = 1.4$), social science ($\mu = 66.6; SE = 2.7$), science ($\mu = 70.9; SE = 2.0$), and synonyms and antonyms ($\mu = 67.6; SE = 2.4$). For each of these academic measures, students with mild intellectual disability earned scores greater than those of students with moderate/severe intellectual disability, but less than those of students with learning disabilities and other high incidence disabilities (refer to Table 3). An ANOVA showed the disability variable to be statistically significant ($p < .001$) for all six academic measures. Bonferroni post-hoc analyses indicated the score for students with mild intellectual disability was statistically significantly different in each of the six measures when compared to the other three disability categorizations. The scores for students with mild intellectual disability on each of the six measures were greater than those for students with moderate/severe intellectual disability, but lower than those of students with the other
high incidence and students with learning disabilities.

**Functional Skills and Self-Determination Measures**

The average score representing all students' mental (i.e., functional) skills was 13.5 (SE = 0.2). Students with mild intellectual disability (μ = 11.0; SE = 0.4) scored slightly below the average, while students with learning disabilities (μ = 14.3; SE = 0.1) and other high incidence disabilities (μ = 14.3; SE = 0.1) earned higher identical scores. Students with moderate/severe intellectual disability (μ = 8.1; SE = 0.4) scored lowest in comparison.

Based on frequency data examining self-determination (i.e., empowerment) assessment scores, statistically significant differences (p < .001) were shown to exist between students with mild intellectual disability and the three comparative groups on the four of the questions (refer to Table 3). Regarding their ability to offer their own opinions and make their own decisions, students with mild intellectual disability scored lower when compared to students with learning disabilities and other high incidence disabilities. When asked whether working hard would lead them to attain what they desire, students with mild intellectual disability scored lower compared to students with learning disabilities, and higher compared to moderate/severe intellectual disability. Finally, when asked whether they believed others honored their choices, students with mild intellectual disability scored significantly higher compared to students with moderate/severe intellectual disability (refer to Table 3).

**Instruction and Instructional Environments**

Students with mild intellectual disability received less instruction (64.4%; SE = 5.3) in life skills classes than students with moderate/severe intellectual disability, but more than students with learning disabilities and other high incidence disabilities (refer to Table 4). Within academic areas, students with mild intellectual disability received less instruction in language arts (89.5%; SE = 3.4), science (64.2%; SE = 5.7), and social studies (73.5%; SE = 4.4) than students with learning disabilities and other high incidence disabilities, but more than students with moderate/severe intellectual disability. In mathematics, students with mild intellectual disability (78.7%; SE = 6.2) received less that students with learning
disability, other high incidence disabilities, and moderate/severe intellectual disability.

When looking at instructional environments, students with mild intellectual disability received the majority of their curriculum instruction in a special education classroom (refer to Table 4). About one-fifth of students with mild intellectual disability received social studies and science in a general education setting; less than one-tenth received language arts and mathematics. By slightly—and consistently—higher levels, students with moderate/severe intellectual disability received most of their curriculum instruction in a special education classroom as well. In contrast, students with learning disabilities and other high incidence disabilities received the majority of their content area instruction in a general education classroom. However, it should be noted a sizable percentage of students with learning disabilities and other high incidence disabilities were also served in special education classrooms for each subject area as well, with the higher frequencies for instruction in a special education classroom occurring for mathematics and language arts.

### Accommodations Received

Of the four disability categories, students with mild intellectual disability received the greatest percentage of the following accommodations: extra time on assignments (72.9%; \(SE = 4.3\)), tests read aloud (53.2%; \(SE = 6.4\)), and modified tests (52.1%; \(SE = 5.1\)) (see Table 5). Students with mild intellectual disability reported receiving the accommodation of extra time on tests (75.2%; \(SE = 4.9\)) at a lower frequency.
percentage than those students with a learning disability and other high incidence disabilities, but significantly more than students with moderate/severe intellectual disability. Students with mild intellectual disability received the accommodations of alternative testing (34.2%; SE = 5.6) and slower paced instruction (52.8%; SE = 6.1) at a lower percentage than students with moderate/severe intellectual disability, but more than students with learning disabilities and other high incidence disabilities (refer to Table 5).

### Discussion

The category of mild intellectual disability experienced a decline in diagnosis and prominence over the past few decades (Bouck, 2007, 2012; Edgar, 1987; Hourcade, 2003). In practice, in place of students being identified as mild intellectual disability, students are being identified with other high incidence disabilities, such as learning disabilities (Beirne-Smith et al., 2006). In research, there is an aggregation of students with mild intellectual disability with students with other high incidence disabilities or students with moderate/severe intellectual disability (Polloway, 2004; 2005; Sabornie et al., 2006). However, results from this study suggest such aggregation or misidentification is not warranted or appropriate, at least at the secondary level. Through a secondary analysis of the NLTS2, researchers identified qualitative and quantitative differences between students with mild intellectual disability and students with moderate/severe intellectual disability, students with learning disabilities and students with other high incidence disabilities—as well as how they are treated.

Across achievement tests, students with mild intellectual disability consistently scored higher than moderate/severe intellectual disability, but lower than students with learning disabilities and other high incidence disabilities; the results supported by statistical analyses. Hence, real average differences exist among secondary students identified with different disabilities on achievement tests. The results from this secondary analysis of national data support those by Sabornie et al. (2006), who suggested students with mild intellectual disability present different cognitive profiles. The differences with regards to achievement data might suggest—or lead one to assume—different instructional practices and other facets of educational programming are warranted.

Trends in academic assessment measures roughly mirrored the type of instructional programming offered to each group. On average, students with mild intellectual disability received less instruction in three of the four academic content areas when compared to students with learning disabilities and students with other high incidence disabilities, but in greater levels than students with moderate/severe intellectual disability. Hence, in practice, students are experiencing different access to academic content based on disability categorization. When considering data on instructional environments, students with intel-

### Table 5

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>MID</th>
<th>M/S ID</th>
<th>LD</th>
<th>Other High Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extra-Time: Tests</td>
<td>75.2% (4.9)</td>
<td>30.5% (8.4)</td>
<td>88.0% (2.0)</td>
<td>87.5% (1.7)</td>
</tr>
<tr>
<td>Tests Read Aloud</td>
<td>53.2% (6.4)</td>
<td>23.5% (6.9)</td>
<td>41.5% (3.8)</td>
<td>38.1% (3.4)</td>
</tr>
<tr>
<td>Modified Tests</td>
<td>52.1% (5.1)</td>
<td>22.9% (6.8)</td>
<td>23.6% (3.0)</td>
<td>23.6% (2.6)</td>
</tr>
<tr>
<td>Alt. Tests</td>
<td>34.2% (5.6)</td>
<td>75.6% (7.0)</td>
<td>13.2% (2.2)</td>
<td>13.8% (1.9)</td>
</tr>
<tr>
<td>Modified Grading</td>
<td>34.3% (5.9)</td>
<td>34.7% (9.5)</td>
<td>9.5% (1.7)</td>
<td>11.3% (1.6)</td>
</tr>
<tr>
<td>Slower Paced Inst.</td>
<td>52.8% (6.1)</td>
<td>78.6% (6.1)</td>
<td>22.8% (2.5)</td>
<td>22.9% (2.2)</td>
</tr>
<tr>
<td>Extra-Time: Assignments</td>
<td>72.9% (4.3)</td>
<td>67.0% (8.9)</td>
<td>60.3% (3.8)</td>
<td>60.8% (3.4)</td>
</tr>
</tbody>
</table>

Note: Other high incidence refers to attention deficit hyperactivity disorder, emotional and behavioral disorder, speech and language disorder.
lectual disability (mild and moderate/severe) were almost entirely served in self-contained special education classrooms, with students with mild intellectual disability participating more in general education classes than moderate/severe intellectual disability but less than students with learning disabilities and other high incidence disabilities. These data also suggest that in practice, educators perceive real differences among students with different disabilities, with suggestions that the least restrictive environment and best classroom to prepare for positive post-school outcomes is different across disability categorizations.

Taken collectively, the findings of this study contest the conclusions offered by Hallahan and Kauffman (1976, 1977) and Gresham et al. (1996) suggesting students with mild intellectual disability and students with other high incidence disabilities (e.g., learning disabilities) possess academic differences that are not educationally relevant with respect to the interventions, accommodations, and classroom placements they receive. Students with mild intellectual disability were—on average—demonstrating achievement levels significantly less than that of their peers identified with other high incidence disabilities, but also significantly greater than their peers identified with moderate/severe intellectual disability. They also, on average, received different educational programming, in terms of curriculum, instructional environments, and accommodations, suggesting there is a presumption in practice that students in these different disability categorizations are also qualitatively and quantitatively different. Hence, with respect to practice, students with mild intellectual disability are likely to benefit from different educational experiences than students with moderate/severe intellectual disability or students with high incidence disabilities, such as students with learning disabilities. In other words, a one-size fit all approach to students with high incidence disabilities or students with intellectual disability is not advised.

Limitations and Future Directions

One limitation of this study is that it is a secondary analysis of the NLTS2, thus making it susceptible to the limitations of the data collection procedures and results of the original study. A second limitation is the use of older data from the NLTS2 (i.e., data from this secondary analysis was collected in 2003–2004). Although data may have changed in the years following data collection from the NLTS2, the findings of this study serve as an indicator of how students with mild intellectual disability are being served, and in which measures they show significant differentiation from students with moderate/severe intellectual disability and students with other high incidence disabilities.

Future research should continue to explore the NLTS2 in terms of secondary analyses to understand the in-school experiences and post-school outcomes of students with mild intellectual disability as well as other disability categories. Further, as it becomes available, researchers need to study the NLTS 2012 to identify more current trends in the placement and services being afforded to students with mild intellectual disability. Related to the implications from this study, future researchers should disaggregate their results for students with mild intellectual disability from students with moderate and/or severe intellectual disability, as well as students with learning disabilities and the aggregate of students with high incidence disabilities. Disaggregation is increasingly important in this era of evidence-based practices; it is important to understand what works and specifically for whom.

References

Bouck, E. C., & Joshi, G. (2012). Functional curriculum and students with mild intellectual disabil-


Utility of Formal Preference Assessments for Individuals Diagnosed with Autism Spectrum Disorder

Justin B. Leaf, Ronald Leaf, Aditt Alcalay, Jeremy A. Leaf, Daniel Ravid, Stephanie Dale, Alyne Kassardjian, Kathleen Tsuji, Mitchell Taubman, John McEachin, and Misty Oppenheim-Leaf

Autism Partnership Foundation and Behavior Therapy and Learning Center

Abstract: The systematic use of reinforcers is an essential component of behavioral intervention for individuals diagnosed with Autism Spectrum Disorder. Today, the use of rigorous formal preference assessments, including paired-preference assessments, are widely conducted to help determine which items to use as reinforcers during intervention. Although paired-preference assessments are widely used there is no experimental evidence whether extensive advance sampling actually produces higher rates of responding compared to in-the-moment analysis of reinforcer effects. The present study compared the rate of responding on a simple sorting task when participants were provided items that were determined as preferred via an extensive paired preference assessment to a teacher selecting items without the use of a paired preference assessment, but rather with an in-the-moment analysis of reinforcer effects. The results indicated no clear difference in the rate of responding, but there were clear differences in terms of efficiency. Clinical implications will be discussed.

Reinforcement can be defined as the presentation of a stimulus causing an increase in the frequency of the targeted behavior in the future (Cooper, Heron, & Heward, 2007). Reinforcement can take many forms, which include: food (e.g., Schreibman, 1975), toys (e.g., Leaf, Sheldon, & Sherman, 2010), praise (e.g., Schreibman, 1975), tokens (e.g., Ayllon & Azrin, 1965), or escape from an undesired event (e.g., Piazza et al., 1997). The provision of reinforcement is widely used to decrease aberrant behaviors and to increase various adaptive behaviors (e.g., Leaf, Dotson, Oppenheim, Sheldon, & Sherman, 2010; Repp & Deitz, 1974). Unfortunately, it is often difficult to identify potential reinforcers for individuals diagnosed with autism spectrum disorders (ASD), which has led to the use of formal preference assessments.

Formal preference assessments are procedures utilized by clinicians to identify which stimuli are preferred by the learner and which stimuli are not preferred, with the presumption that the preferred stimuli are more likely to function as potential reinforcers during teaching. There have been several variations of preference assessments that have been utilized to identify reinforcers, including: interviews (e.g., Piazza, Fisher, Hagopian, Bowman, & Toole, 1996), single stimulus approaches (e.g., Green et al., 1988), paired preference assessments (e.g., Fisher et al., 1992), multiple-stimulus without replacement (e.g., Restar & Noell, 2008), and multiple stimulus with replacement (e.g., Leon & Iwata, 1996). Results from the studies on these various preference assessment procedures have shown that a learner’s preference identified from formal preference assessments is highly correlated with that item’s effectiveness as a reinforcer. Formal preference assessments have been utilized for a wide variety of populations ranging from typically developing children to children diagnosed with ASD (Leaf et al., 2012; Restar & Noell, 2008).

One type of preference assessment commonly evaluated and implemented is a paired-preference assessment (e.g., Fisher et al., 1992). A paired preference assessment consists of the following components. First, the teacher identifies (e.g., via interviews) several possible reinforcing items (e.g., toys, social
reinforcers, and edibles). Next, the teacher presents the first pair of items (i.e., pairs item one against item two) and has the learner indicate which item he or she wants to interact with or consume. Third, the learner then has an opportunity to play with or consume the item that he or she selected. Pairings continue until every item is paired against all other items. Based upon the learner’s selections, the teacher creates a hierarchy from the learner’s most preferred items to least preferred items. The most preferred items are then used as reinforcers during teaching opportunities.

Fisher and colleagues (1992) were the first to empirically validate the paired preference procedure. In this study, the researchers compared the paired-preference assessment to a single stimulus preference assessment for four participants diagnosed with intellectual, physical, and developmental disabilities. The researchers utilized a concurrent operant paradigm to compare the two preference procedures. Results of the study indicated that, although both procedures were accurately able to differentiate highly preferred to less preferred items for the participants, the paired preference assessment was able to show greater differentiation amongst items. Since Fisher’s original study, the paired preference assessment has been used to identify highly preferred and less preferred items for a wide variety of populations, including high functioning individuals diagnosed with ASD (e.g., Leaf et al., 2012).

Windsor, Piche, and Locke (1994) compared a paired preference method, a group preference method, and a staff ranking method to identify preferences amongst food and drink stimuli for eight adults with profound developmental disabilities. Results of the study showed that the group presentation and paired preference assessment yielded similar results. However, the results indicated that staff rankings did not match the paired preference assessment in differentiating between the different stimuli. In a more recent study, Cote and colleagues (2007) compared teacher report to a paired-preference assessment in identifying and determining the reinforcement value of various stimuli for typically developing toddlers. The results of this study indicated that items identified as more preferred in the paired preference assessment were more reinforcing than the items identified through teacher rankings; however, the items selected through both procedures served as reinforcers.

The research on paired preference assessments has resulted in wide spread clinical use. Graff and Karsten (2012) surveyed 406 professionals, 32% of whom were Board Certified Behavior Analysts (BCBA) or Board Certified Assistant Behavior Analysts (BCaBA), on their use of a wide variety of formal preference procedures. Fifty two percent of all respondents reported using at least one direct formal preference assessment; 89% of BCBAs or BCaBAs reported using at least one direct preference assessment throughout the course of intervention. Furthermore, 36% of all respondents and 70% of BCBAs/BCaBAs have reported previously using a paired-preference assessment. Finally, Graff and Karsten reported that approximately 45% of BCBAs use a preference assessment at least on a monthly basis; ranging from 3% using formal preference assessments several times a day, 18% using formal preference assessments at least once a month, and 24% using formal preference assessments somewhere in between.

Previous research has shown that paired preferences are: (a) effective in identifying highly preferred and less preferred stimuli, which are highly correlated with reinforcement value; (b) able to differentiate the relative preference value of various stimuli across different populations; (c) more accurate than staff interviews or surveys; and (d) widely implemented clinically. What is not known, however, is whether paired preference assessments are actually necessary in changing the behavior for individuals diagnosed with ASD or if there is an alternative procedure that can produce similar effects.

In-the-moment reinforcer analysis requires a teacher to select which item is best suited as a reinforcer based upon current circumstances, instead of selecting an item based upon (prior) formal preference assessments. In-the-moment reinforcer analysis takes into account: (a) analyzing learner affect (e.g., facial expressions, verbal statements, body language); (b) analyzing how the learner interacts with the item; (c) how often the teacher utilized the item during previous sessions or trials; (d) the learner’s motivation to gain ac-
cess based upon responding; (e) identifying features of other preferred items and incorporating those features into the new item; and (f) the teacher’s ability to condition alternative items as reinforcers as described by Leaf et al. (2012). In the 2012 study, Leaf and colleagues utilized a similar in-the-moment reinforcer analysis strategy combined with an observational conditioning procedure to increase a participant’s preference for items that were originally non-preferred.

The purpose of the present study was to evaluate the utility of a formal paired preference assessment as compared to in-the-moment reinforcer analysis of reinforcer effects in maintaining or increasing the rate of responding during a simple sorting task for three children diagnosed with ASD. Specifically, we evaluated whether there was a difference in responding under the two conditions and the total time the participant spent in each condition.

Method

Participants and Setting

Wyatt was a four-year-old boy independently diagnosed with Autistic Disorder. Wyatt had a Wechsler Preschool and Primary Scale of Intelligence-IV (WPPSI-IV) full scale IQ score of 86, a Vineland Adaptive Behavior Scales (VABS) composite score of 83, a Gilliam Autism Rating Scale-2 (GARS-2) Autism Quotient of 100 (probability of autism: very likely), and a Social Skills Improvement System-Parent Version (SSIS-P) standard score of 92. Wyatt was conversational and displayed beginning play skills (e.g., imaginative play and parallel play). Wyatt also displayed several stereotypic behaviors, including: twirling items, lining up items, stacking items, and making perseverative statements. Prior to the study, Wyatt had received seven months of intensive early behavioral intervention.

Henry was a five-year-old boy independently diagnosed with Autistic Disorder. Henry had a WPPSI-IV full scale IQ score of 85, a VABS composite score of 72, GARS-2 Autism Quotient of 87 (probability of autism: very likely), a SSIS-P standard score of 88, and a Autism Diagnostic Observation Schedule (ADOS) total score of 19 (consistent with Autistic Disorder). Henry was conversational and displayed advanced play skills (e.g., imaginative play, parallel play, cooperative play, and game play). Henry displayed a few stereotypic behaviors, including perseveration on topics of special interest. Prior to the study, Henry had received 14 months of intensive early behavioral intervention.

Marty was a four-year-old boy independently diagnosed with Autistic Disorder. Marty had a WPPSI-IV full scale IQ score of 93, a VABS composite score of 81, GARS-2 Autism Quotient of 78 (probability of autism: possible), and a SSIS-P standard score of 78. Marty was conversational and displayed advanced play skills (e.g., imaginative play, parallel play, cooperative play, and game play). Marty displayed stereotypic behaviors, including perseverations and rigidity. Additionally, Marty displayed non-compliance, tantrums, and aggression towards staff members. Prior to the study, Marty had received two years of early intensive behavioral intervention.

For Wyatt and Henry, research sessions took place in a small research room located at the private agency that provides behavioral intervention for individuals diagnosed with ASD. For Marty, research sessions took place in the same research room and in his home therapy room.

Teachers

The study was conducted by teachers who were employed for at least seven months by a private agency that provides intensive behavioral treatment for individuals with ASD. Each teacher had a history with each of the participants. All teachers were trained in the principles of applied behavior analysis, reinforcement, and ASD. Each teacher was randomly assigned to either the preference assessment condition (described below) or the in-the-moment reinforcer analysis condition (described below). Teachers were kept blind from each other’s findings, so that they could not discuss the results of a session or the reinforcers utilized within each condition. Therefore, the teachers assigned to the in-the-moment reinforcer analysis did not know the results of the paired preference assessment. If a teacher was randomly assigned to the preference assessment condition, he or she was responsible for
running the two paired preference assessments and running all preference assessment condition sessions (described below). If a teacher was randomly assigned to the in-the-moment reinforcer analysis, he or she was responsible for running the in-the-moment reinforcer analysis (described below). Teachers assigned to either condition were able to run the control condition (described below).

**Pre-Baseline**

**Ranking Reinforcers.** To determine which potential reinforcers would be utilized in the reinforcement assessment (see below) the researcher (first author) emailed each member of the participant’s clinical team (i.e., teachers and program supervisors) requesting each member to provide a list of all items (e.g., light toys, bouncy balls, figurines, etc.) or social activities (e.g., red light-green light game, tickles, or high fives) that the participant enjoyed. The researcher asked that each member respond via email and that each member of the clinical team not discuss the list with other members of the team. Once the researcher received each member’s list, the researcher created a larger list consisting of all the items that were identified by any member of the clinical team.

Next, the researcher emailed each member of the participant’s clinical team the list of all of the reinforcers and asked each person to numerically rank the list from the participant’s presumed most preferred item to the participant’s presumed least preferred item. Once again, the researcher asked that the members of the clinical team not discuss their rankings with each other. Once the researcher received all of the rankings, he averaged the rankings across all members of the clinical team. The top seven presumed preferred items from this ranking were used in the paired preference assessment (described below) and in the in-the-moment reinforcer analysis (described below).

**Paired Preference Assessments.** Prior to the reinforcement assessment, two paired-choice preference assessments (Fisher et al., 1992), across two different days, were conducted in order to identify which three items were to be used as reinforcers in the preference assessment condition (see below). The teachers assigned to the preference assessment condition evaluated a total of 10 items; seven items were selected based upon staff rankings (described above) and three were not identified by the rankings but were selected because each item was an age appropriate toy or activity for the participant.

Prior to starting the first preference assessment, the participants were given the opportunity to play with each item for approximately 30 s. Next, the teacher conducted the paired preference assessment. The paired preference assessment consisted of the teacher holding out two items (one in each hand), asking the participant to select an item to play with, and having the participant physically touch that item. After the participant physically selected the item, the participant was provided with 30 s access to the item. During this time the teacher socially engaged with the participant and the item. This was continued until all items were compared against each other item one time.

The second day of the paired preference assessment was identical to the first day with the exception that the participant did not have 30 s access to each item before the assessment. The top three items that were selected most frequently across the two days were used as the reinforcers in the preference assessment condition and all ten items were used in the in-the-moment reinforcer analysis condition. The top three preferred items identified for Wyatt were two big toy cars, a light spinner, and four small toy cars. The top three preferred items identified for Henry were a Magnadoodle, Hex Bug, and Batman and Joker figurine. The top three preferred items identified for Marty were a Imaginix Dinosaur, Angry Bird stuffed animals, and two light sabers. Participants only had access to these items during research sessions, to minimize satiation effects.

**General Procedure**

Research sessions were conducted three days per week for Wyatt and Marty and four days per week for Henry. A single research session was conducted per day, which consisted of the preference assessment condition, in-the-moment reinforcer analysis condition, and control condition. The order of each of these
conditions was randomized and counterbalanced. Participants received at least a 10 min break in between each of the conditions.

The general procedure across all three conditions was as follows. First, the teacher would bring the participant in the room and place a color mat that corresponded with one of the conditions (i.e., a red mat for the in-the-moment reinforcer analysis condition, a blue mat for the preference assessment condition, and a white mat for the control condition) on the table. Next, the teacher placed one large bin containing 450 poker chips (i.e., 150 red chips, 150 green chips, and 150 white chips) on the table and placed three smaller bins next to the larger bin. The three smaller bins each had a single color poker chip placed at the bottom of the bin; one bin had a red poker chip, one bin had a green poker chip, and one bin had a white poker chip placed at the bottom.

The teacher then called the participant over to the table and had the participant sit directly across from him/her. The teacher then provided the following instruction: “I want you to sort the chips (while pointing to the large bin), the red chips go in here (pointing to the bin with the red chip placed in it), the green chips go in here (pointing to the bin with the green chip placed in it), and the white chips go in here (pointing to the bin with the white chip placed in it). You can do as many as you want, but you do not have to do any if you do not want to. You will be working for toys this time (during the preference and in-the-moment reinforcer analysis conditions) or there will be no toys this time (during baseline or control condition). Ready, set, go!” The teacher then started an electronic timer and gave the participant one-minute to sort as many chips as he chose to sort. The teacher provided no prompts, did not redirect any off-task behavior, provided no reinforcement, nor any social interaction during the one-minute period of time. The participant was allowed to sort as many chips as he chose; if the participant engaged in off-task behavior (e.g., falling to the floor), he was allowed to do so.

Once the timer rang, the teacher told the participant to “stop” and blocked the participant from placing any additional chips in the smaller baskets. Next, the teacher counted out loud the number of chips that were correctly sorted in each of the buckets, while placing those chips back into the larger bin. Finally, a consequence was provided dependent upon the condition and whether the participant reached the targeted number (described below). There were a total of six trials per condition per session.

Each day the researcher informed the teachers the number of chips that the participant had to sort in order to receive reinforcement in the preference assessment condition and the in-the-moment reinforcer analysis condition. This number was never revealed to the participants. During the first session of intervention, in order to earn reinforcement, participants were required to sort 20% more chips than their average during the baseline session. After every two consecutive sessions in which the participant reached the targeted number of chips, in either the preference assessment condition or the in-the-moment reinforcer analysis condition, the targeted number of chips was increased by 20%. If, after three consecutive sessions, the participant was unable to meet the targeted number in any of the conditions, then the researcher reduced the targeted number by 20%.

**Conditions**

**Baseline.** A single baseline session was conducted to determine the initial average rate of chip sorting per trial for each participant. This session consisted of six trials as described above. At the end of each trial the teacher thanked the participant regardless of the amount of chips he sorted correctly. No further reinforcement was provided during baseline.

**Preference Assessment Condition.** Each preference assessment condition session consisted of six total trials. The teachers who were randomly assigned to this condition ran every research session. Each trial was run using the procedure described above. If the participant sorted enough chips to reach or exceed the target number, the teacher provided him with one-minute access to one of the three items or activities. The three top items selected in the paired preference assessment were evenly dispersed across the six trials; the participant had an opportunity to earn each item two times.
per session. The order of delivery of the reinforcers was predetermined prior to each session. Only one item was provided to the participant per trial. If the participant sorted enough chips they were told “You got enough, we can play with . . .”; if the participant did not sort enough chips to reach or exceed the target number, the teacher told the participant, “You did not sort enough chips and there will be no toy this time.”

**In-The-Moment Reinforcer Analysis Condition.** Each in-the-moment reinforcer analysis condition session consisted of six total trials. The teachers who were randomly assigned to this condition ran every research session. Each trial was run using the procedures described above. If the participant sorted enough chips to reach or exceed the target number, the teacher provided him with one-minute access to any of the ten items that were assessed in the paired preference assessment. The teacher had the freedom to select any of the items at his or her own discretion. For example, the teacher could select the same item several times throughout the session or could choose a different item each trial. Only one item was selected per trial. If the participant sorted enough chips, they were told “You got enough, we can play with . . .”; if the participant did not sort enough chips to reach or exceed the target number, the teacher told the participant, “You did not sort enough chips and there will be no toy this time.”

**Control Condition.** Both teachers assigned to the preference assessment condition and the in-the-moment reinforcer analysis condition ran control sessions. Control sessions consisted of six trials as described above. Participants were made aware that they would not be working for items during this condition (described above). At the end of each trial the teacher thanked the participant, regardless of the amount of chips he sorted correctly, and immediately began the next trial. No further reinforcement was provided during control condition sessions.

**Dependent Variable and Measures**

There were three dependent variables evaluated in this study. The primary dependent variable was the average rate of responding per trial in each condition. At the end of each session, for each condition, the researchers calculated the total number of chips the participant sorted in the session and divided that number by the total number of trials (n = 6).

The second dependent variable was an efficiency measure. This was determined by the total amount of time the participant spent in each of the three conditions. For the preference assessment condition the total time included both the time the participants spent in the preference assessment condition and the time the participants spent in the two paired-preference assessments.

The third measure evaluated was the stimuli selected by the teacher in the in-the-moment reinforcer analysis condition. The researchers evaluated the percentage of times the teacher selected each of the ten stimuli across all in-the-moment reinforcer analysis sessions.

**Design**

An alternating treatments design was used to evaluate the effects of the three conditions (described above) on chips sorted per session. The sequence of the conditions each research session was randomly determined ahead of time and counterbalanced to control for sequence effects.

**IOA and Treatment Fidelity**

The teacher running the session scored the participant’s responses during every session. A second observer (first author) simultaneously and independently recorded participant responses during 100% of baseline sessions, 48% of preference assessment condition sessions, 68% of in-the-moment reinforcer analysis sessions, and 56% of control condition sessions. Inter-observer agreement was calculated by totaling the number of agreements (i.e., number of trials in which both observers scored the same amount of chips sorted) divided by the number of agreements plus disagreements and converting this ratio to a percentage. IOA was 100%.

A second observer (first author) independently scored whether the teacher assigned to the preference assessment condition and the teacher assigned to the in-the-moment analysis condition, and the teachers implementing the control condition implemented correct
instructor behaviors in 48%, 68%, and 56% of sessions, respectively. Correct instructor behaviors per trial consisted of: (a) placing all four bins on the table; (b) reading the script (described above) correctly; (c) setting a timer for one-minute and allowing the participant to sort the chips for one-minute; and (d) providing the correct reinforcer or providing no reinforcer in the control condition. Treatment fidelity was 100%.

Results

Rate of Responding

The primary dependent variable was each participant’s responding during the three conditions. The results of this measure are depicted in Figure 1 and Figure 2. For Figure 1, each panel depicts a different participant’s responding. Across the x-axis is the number of sessions and across the y-axis is the average number of chips sorted per trial within a given session, across the three conditions. It must be noted that the range along the y-axis varies from panel to panel, due to the amount of chips that each participant sorted. Participant responding during the single baseline session is not represented on the graph, but rather reported in the text below. For Figure 2, the x-axis represents each of the three participants and the y-axis depicts the average number of chips sorted per trial throughout all sessions.

Wyatt’s data per session is reported in the top panel of Figure 1 and overall responding on Figure 2. During the baseline session, Wyatt sorted an average of 7.3 chips per trial. Following the baseline session, Wyatt participated in five research sessions; Wyatt’s average rate of responding was 10.09, 10.34, and 1.77 chips sorted for the preference assessment, in-the-moment reinforcer analysis, and control condition, respectively (see Figure 2). This represents a 38% and 41% increase from baseline levels for the preference assessment condition and in-the-moment reinforcer analysis condition, respectively. Results show that there was a slightly stronger reinforcement effect in the in-the-moment reinforcer analysis condition. Additionally, a clear extinction effect was observed during the control condition; his rate of responding steadily and quickly decreased.

Henry’s data per session is reported in the middle panel of Figure 1 and overall responding on Figure 2. During the baseline session, Henry sorted an average of 28.1 chips per trial. Following the baseline session, Henry participated in 13 research sessions; Henry’s average rate of responding was 25.46, 25.57, and 23.9 for the preference assessment condition, in-the-moment reinforcer analysis condition, and control condition, respectively (see Figure 2). Although this does not represent an overall increase from baseline, Henry’s data was variable; on some days an increase was seen and on other days a decrease was demonstrated (see Figure 1). Therefore, for Henry, there was no discernible effect on the rate of responding in either of the reinforcement conditions.

Marty’s data per session is reported in the bottom panel of Figure 1 and overall responding on Figure 2. During the baseline session, Marty sorted an average of 10.3 chips per trial. Following the baseline session, Marty participated in 7 research sessions; Marty’s average rate of responding was 17.11, 18.44, and 9.53 for the preference assessment condition, in-the-moment reinforcer analysis condition, and control condition, respectively (see Figure 2). This represents a 66% and 79% percent increase from baseline levels for the preference assessment condition and in-the-moment reinforcer analysis condition, respectively. Results show that there was a slightly stronger reinforcement effect in the in-the-moment reinforcer analysis condition. Additionally, a clear extinction effect was observed during the control condition; his rate of responding steadily and quickly decreased.

Efficiency Measure

Table 1 displays the efficiency results across the three participants. Efficiency was measured by the total amount of time a participant spent in each condition; for the preference condition, efficiency included the total amount of time spent in the paired preference assessments. Wyatt spent approximately one-minute longer in the in-the-moment reinforcer analysis condition than the
preference assessment condition. However, when including the time required to conduct the two paired preference assessments, Wyatt spent approximately 85 minutes longer in the preference assessment condition to achieve similar results. Henry spent approximately 13 minutes longer in the in-the-moment reinforcer analysis condition than the preference assessment condition. However, including time for the two paired preference assessments, Henry spent approximately 78 minutes longer in the preference assessment condition.
to achieve similar results. Marty spent approximately one minute longer in the in-the-moment reinforcer analysis condition than the preference condition. However, with time for the two paired preference assessments included, Marty spent approximately 85 minutes longer in the preference assessment condition to achieve similar results.

Distribution of Stimulus Selection

Figure 3 depicts the percentage of trials that each of the 10 reinforcers was delivered by the teacher during trials in which the participant reached or exceeded the target number in the in-the-moment reinforcer analysis condition. Along the x-axis are the ten different toys that the teacher could have selected; the items are placed in order of preference based upon the results of the paired preference assessments, with the most preferred reinforcer being closer to the y-axis and the least preferred item being farthest away from the y-axis. Along the y-axis is the percentage of selections for each of the items during trials in which the participant responded rapidly enough to earn a reinforcer. Each panel represents the results for a different participant.

For Wyatt (top panel), the two items that were delivered most frequently were two big toy cars (22.2%) and a toy pirate ship (22.2%). The big toy cars was the item that
was deemed most preferred during the paired preference assessment; the toy pirate ship was the third least preferred. The other two items that were determined most preferred and used as reinforcers in the preference condition were a light spinner and four little cars; the teachers in the in-the-moment reinforcer analysis condition selected these items during 5.5% and 5.5% of opportunities, respectively. Overall, the teacher in the in-the-moment reinforcer analysis condition selected items that were not used in the preference assessment condition more frequently than items that were used in the preference assessment condition.

For Henry (middle panel), the two items that were delivered most frequently were a magnadoodle (33.3%) and a whoopee cushion (25.7%). The magnadoodle was the item that was deemed most preferred during the paired preference assessment; the whoopee cushion was the fourth most preferred. The other two items that were determined most preferred and used as reinforcers in the preference assessment condition were a hex bug and a batman and joker figurine; the teachers in the in-the-moment reinforcer analysis condition selected these items during 7.6% and 0% of opportunities, respectively. Overall, the teacher in the in-the-moment reinforcer analysis condition selected items that were not used in the preference assessment condition more frequently than items that were used in the preference assessment condition.

For Marty (middle panel), the three items that were delivered most frequently were the angry bird stuffed animals (29.1%), light sabers (20.8%), and a hungry hippo game (20.8%); the angry bird stuffed animals and the light sabers were two of the most preferred items selected during the paired preference assessment; the hungry hippo game was the least preferred item selected during the paired preference assessment. Overall, the teacher in the in-the-moment reinforcer analysis condition selected items that were used in the preference assessment condition more frequently than items that were not used in the preference assessment condition; however, the teacher selected toys not used during the preference assessment condition during 37.6% of opportunities.

**Discussion**

The purpose of this study was to determine if there would be a differential rate of responding for three young children diagnosed with ASD when a teacher used the top three toys identified by a formal paired preference assessment, versus a teacher using in-the-moment reinforcer analysis to select amongst ten items, versus a control condition (no external reinforcement). Results of this study showed that for two participants (Wyatt and Marty) there was a slightly higher rate of responding in the in-the-moment reinforcer analysis condition as compared to the preference assessment condition, and both reinforcement conditions were considerably superior to the control (extinction) condition. Thus, for these two participants, external reinforcement

**TABLE 1**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Condition</th>
<th>Paired Preference Assessment Time</th>
<th>Session Time</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wyatt</td>
<td>Preference</td>
<td>86 Min 11 Sec</td>
<td>78 Min 8 Sec</td>
<td>164 Min 19 Sec</td>
</tr>
<tr>
<td></td>
<td>In-the-moment Reinforcer Analysis Condition</td>
<td>0 Min 0 Sec</td>
<td>79 Min 2 Sec</td>
<td>79 Min 2 Sec</td>
</tr>
<tr>
<td>Hank</td>
<td>Preference</td>
<td>91 Min 14 Sec</td>
<td>207 Min 37 Sec</td>
<td>298 Min 51 Sec</td>
</tr>
<tr>
<td></td>
<td>In-the-moment Reinforcer Analysis Condition</td>
<td>0 Min 0 Sec</td>
<td>220 Min 17 Sec</td>
<td>220 Min 17 Sec</td>
</tr>
<tr>
<td>Marty</td>
<td>Preference</td>
<td>85 Min 21 Sec</td>
<td>115 Min 47 Sec</td>
<td>201 Min 8 Sec</td>
</tr>
<tr>
<td></td>
<td>In-the-moment Reinforcer Analysis Condition</td>
<td>0 Min 0 Sec</td>
<td>116 Min 26 Sec</td>
<td>116 Min 26 Sec</td>
</tr>
<tr>
<td></td>
<td>Condition</td>
<td>0 Min 0 Sec</td>
<td>74 Min 12 Sec</td>
<td>74 Min 12 Sec</td>
</tr>
</tbody>
</table>
Figure 3. In-The-Moment Reinforcer Analysis Teacher Percentage of Selection across All Items and All Sessions.
was necessary to evoke high rates of responding. For the third participant (Henry), there was little difference in the rate of responding among the three conditions indicating that external reinforcement was not necessary to produce a high rate of responding. Finally, there was a considerable degree of overlap in reinforcer selection between teachers assigned to the in-the-moment reinforcer analysis and teachers assigned to the preference assessment conditions; therefore, it is not surprising to see similar rates of responding across the two conditions. However, the teachers assigned to the in-the-moment reinforcer analysis condition were able to select these reinforcers without the use of formal assessments.

Secondly, the efficiency results show that conducting paired preference assessments significantly adds to the total amount of teacher and learner time spent in the preference assessment condition in order to achieve a high rate of responding. The efficiency results also show that selecting a range of items, based on in-the-moment reinforcer assessment can produce responding that is equal to selecting only items deemed highly preferred via paired preference assessments. Therefore, the results for this study showed that conducting paired preference assessments provided no real advantage in performance level for these three participants, and carried the disadvantage of taking more time.

Previous researchers have shown that formal preference assessments result in better identification and differentiation amongst several items and are more accurate in finding effective reinforcers as compared to interviews or staff opinions (e.g., Fisher et al., 1992). However, most of these studies compare the single item identified as most preferred via the formal preference assessment to the single item identified as most preferred via a staff interview or questionnaire, rather than using the top tier items (e.g., top three) identified as highly preferred via a formal preference assessment.

The other procedural difference in the present study is the use of the entire range of items identified as reinforcers via interviews and allowing the teacher to use discretion at each occurrence of reinforcement delivery versus the use of only the single item that was most frequently reported to be reinforcing based on interviews. While it is possible that the formal preference assessment condition in the present study may have been disadvantaged by the inclusion of the second-most and third-most preferred items, thereby reducing the overall average preference level, it is also possible that the use of three items was a benefit in mitigating satiation effects; further research is necessary to answer this question. Based on the procedures used in the present study, however, the similar responding rates across the two conditions and three participants, still clearly indicates that paired preference assessments provided no advantage over the in-the-moment reinforcement analysis, and the time spent conducting paired preference assessments may be detrimental to the participants.

The average amount of time to complete a single paired preference assessment was approximately 44 min. Although this time is not significant in length, it can be significant if a preference assessment is provided on a more frequent (e.g., several times a day or every day) basis. According to Graff and Karsten (2012), 3.8% of all BCBAs or BCaBAs reported that they implement formal preference assessments on a daily basis. Based upon our average amount of time spent (44 min) conducting a paired preference assessment, conducting one on a daily basis would result in approximately 3 hours and 40 min per week, or 190 hours and 40 min per year (approximately 9 weeks of a 20 hour behavioral program), of time spent just in formal preference assessments. If a clinician were to implement a paired preference assessment on a weekly basis, it would still result in approximately 38 hours and 8 min per year spent conducting paired preference assessments.

Given that there is no significant difference in the rate of responding, the large amount of time an individual may spend conducting preference assessments, and that previous research has shown that preferences can be switched from a highly preferred item to a less preferred item when a therapist uses conditioning procedures (Leaf et al., 2012), it would appear that the formal use of preference assessments is unjustified and that time would be better spent elsewhere.

This study is not without its limitations,
which precludes us from making global confirmative statements about the use of formal preference assessments. For one, the study only included individuals diagnosed with ASD who would be considered higher functioning. Although this population may not need the explicit use of formal preference assessments, researchers have conducted studies evaluating formal preference assessments for individuals who are typically developing (e.g., Restar & Noell, 2008) and, based upon the survey results from Graff and Karsten (2012), it can be presumed that formal preference assessments are being clinically implemented with similar populations. Anecdotally, that is certainly consistent with our observations. Nevertheless, future researchers should compare using items selected based upon a paired preference assessment to items selected based upon in-the-moment reinforcer analysis for a wider range of individuals diagnosed with ASD.

Second, this study only evaluated the utility of preference assessments as it applies to a simple rate response measure. Therefore, it is not known if differences would be observed for responding on more difficult rate tasks, on other learning tasks (e.g., receptive labeling or expressive labeling), or for decreasing aberrant behaviors. Third, Henry responded at equivalent rates for all three conditions thereby weakening the demonstration of functional control of his responding, since he continued to respond at a high rate during the control condition (extinction). Fourth, we only utilized a paired preference assessment, so it is not known what the results would have been if we had utilized other formal assessments. Finally, the research included teachers who were well trained in the principles of applied behavior analysis, ASD, and reinforcement. Therefore, it is not known what the effects would be using a novice teacher. Although the use of formal preference assessments may help novice teachers identify and utilize reinforcers more effectively, this should not preclude researchers and clinicians from finding effective ways to train teachers in better identifying, creating, and utilizing reinforcement.

Despite these limitations, results of the study showed that, for our three participants, there was no clear difference in the rate of responding; however, the amount of time a participant spent during the preference assessment condition was substantially longer than time spent in the in-the-moment reinforcer analysis condition. Therefore, the results of this study indicate that the use of paired preference assessments was not warranted for our three participants. Future researchers should expand the demonstration of clinical utility and effectiveness of reinforcers and take into account the time penalty required for conducting extensive preference assessments. Additionally, the in-the-moment reinforcer analysis could allow the learner to choose which of several reinforcers he or she would like to access on each occasion of reinforcer delivery.

References


---

Received: 9 January 2014  
Initial Acceptance: 14 March 2014  
Final Acceptance: 15 June 2014
Use of Say-Do Correspondence Training to Increase Generalization of Social Interaction Skills at Recess for Children with Autism Spectrum Disorder

Nancy Rosenberg, Marissa Congdon, and Ilene Schwartz
University of Washington
Debra Kamps
University of Kansas

Abstract: Research suggests that while social skills groups in school settings can be effective for students with Autism Spectrum Disorder (ASD), generalization of the skills and behaviors learned in these groups to other settings can be problematic. This study assessed the use of a say-do correspondence intervention to increase generalization at recess of social interactions skills previously learned in a social skills group for students with ASD. The participants were three first graders who had participated in intensive social skills instruction for over a year but who were not generalizing their acquired skills to recess. The say-do correspondence involved the participants identifying before recess who they were going to talk to at recess and then receiving access to reinforcers after recess if they had talked to the student they had identified. A multiple-baseline across participants design was used to assess the impact of the intervention on the number of social exchanges between the child with ASD and other children during recess. Results showed that the number of social exchanges increased for all participants. Implications for practice in public school settings are discussed.

As increasing numbers of children with Autism Spectrum Disorder (ASD) enter public schools, school staff are struggling to understand how to teach these children the social skills necessary to be successful in the classroom, playground, and beyond. The need for interventions is clear: problems in the social area are a core deficit of ASD and a part of the diagnostic criteria. But how to successfully target these deficits in the school environment with effective interventions that can be implemented by school personnel is less clear.

One approach to addressing these children’s need for explicit social skills instruction in schools is to conduct social skills groups that meet on a regular basis and systematically address different aspects of social interaction. While numerous studies have demonstrated the efficacy of these groups in building social competency (e.g., Lopata, Thomeer, Volker, Nida, & Lee, 2008; Owens, Granader, Humphrey, & Baron-Cohen 2008; Yang, Schaller, Huang, Wang, & Tsai, 2003), several reviews have concluded that a major weakness of these groups appears to be generalization: participating children become adept at demonstrating the skills within the social skills group but do not demonstrate the skills in other school environments or settings (Bellini, Peters, Benner, & Hopf, 2007; White, Keonig, & Scahill, 2007). The research suggests that children may need to have at least some intervention directly within the generalization environments (that is, the environments in which the children actually need to demonstrate these skills) if the children are going to demonstrate the skills learned in social skills groups in these environments.

Recess can be both an ideal and a challenging time to work on social skills. Recess has been found to be important for physical, cog-
nitive, and social development (Lang et al., 2011). It is ideal for social development in that it provides ample opportunities for natural, playful interactions with a wide variety of typical peers. It is also one of the few times during the school day that children can engage in child-directed activities, choosing where and what they want to play, and with whom they want to play. With in-class time increasingly focused on academics prescribed by a centralized curriculum, it can often be difficult to find appropriate opportunities for students with ASD to practice the social skills they need to acquire outside of designated social skills groups, and recess may be one of the few available times.

Recess, however, can be challenging both for students with ASD and for school staff. For the student with ASD, recess can be unstructured and chaotic, making it a difficult time for students who generally do best in structured, predictable situations. For school personnel, staffing issues at recess can be challenging (Koegel, Kuriakose, Singh, & Koegel, 2012). Many times, schools provide minimal staffing at recess, typically just enough staff to oversee student safety. Recess is often a time that staff takes breaks or utilizes planning time. Rarely is there enough staff to provide intensive 1:1 intervention for a student. In addition, some school personnel do not believe that recess time is instructional time. They believe that recess should be considered a break for students as well as for teaching staff and are thus reluctant to present what they perceive as intrusive demands during this time.

Research suggests that because of these factors and probably others, children with ASD are not being socially successful at recess. In a review of recess interventions, Lang et al. (2011) reported that baseline levels across studies showed that students with ASD tended to demonstrate high levels of stereotypy and low levels of social interactions at recess. Baumerger, Shulman, and Agam (2003) found that when compared to typical children, children with ASD interact with other children significantly less often at recess.

Even when adults are both available and willing to support students at recess, direct intervention to prompt and assist a child with ASD in social interactions in natural school environments such as recess can be problematic. The moment an adult intervenes in a natural child to child interaction, the nature of that interaction is affected (Giangreco, Edelman, Luiselli, & MacFarland, 1997; Malmgren & Causton-Theoharis, 2006). It is likely that the behavior of all the children involved in the interaction, not just the behavior of the child with ASD, is influenced by the adult presence.

In summary, strategies for helping children with ASD to generalize social skills across settings and activities at school, and increase frequency of interactions with other children at recess are needed. These interventions, however, must take into account the fact that staffing is likely to be sparse at recess and that strategies that involve direct adult prompting or other direct intervention are likely to detract from the natural interactions between children and may feel inappropriate to some adults. Methods to promote the use of learned social skills in these natural environments without the necessity of direct adult intervention into the ongoing interactions are needed.

The use of a say-do correspondence intervention could be ideal in such situations. Say-do correspondence traditionally involves the modification of nonverbal behavior via changes in verbal behavior. In say-do correspondence (also called correspondence training), a person receives access to a reinforcer for demonstrating correspondence between “saying:” stating what he is going to do at some point in the future, and “doing:” actually doing what he said he was going to do. In other words, the behavior that is reinforced is making a plan and then implementing that plan. Research suggests that after a period of reinforcement of correspondence between saying and doing, change can be enacted in nonverbal behavior (the “do”) via control of the verbal statement (the “say”). A substantial body of research has established the effectiveness of correspondence training for persons with and without disabilities and for both adults and children (for a review, see Bevill-Davis, Clees, & Gast (2004). In the first study of say-do correspondence, Risley and Hart (1968) increased the use of specific play materials by preschoolers, establishing correspondence between a stated intention by a child to play with specific materials and then actually playing with the specific play materials. In an example
of say-do correspondence with children with disabilities. Whitman, Scibak, Butler, Richter, and Johnson (1982) reported on three experiments using say-do correspondence to decrease out-of-seat behavior, increase appropriate sitting posture, and increase on-task behavior for children with intellectual disability. Their third experiment was unique in that it attempted to increase on-task behavior for three non-verbal students using a non-verbal “say” component, requiring the students to “show” the intended behavior rather than state it verbally. On-task behavior increased for all three students.

Similarly, it is worth noting that while traditionally, the “do” portion of say-do correspondence is defined as a nonverbal behavior; it does not actually have to be nonverbal. Rogers-Warren and Baer (1976), for example, used say-do correspondence to increase a verbal behavior: the amount of praise directed towards peers; similarly, Guevremont, Osnes, and Stokes (1986a) used correspondence training to increase peer-directed talk for a four-year-old child. Thus, the traditional definition of say-do correspondence as correspondence between a verbal and a non-verbal behavior is probably over specific; both the say and the do components can be either verbal or non-verbal behaviors.

Correspondence training has often been cited as being ideal in situations where an instructor has ready accessibility to the individual during the “say” component of the intervention but wishes the individual to demonstrate the nonverbal behavior in situations without direct monitoring or reinforcement (e.g., Guevremont, Osnes, Stokes, 1986a, b; Morrison, Sainato, Benchaabane, & Endo, 2002; Whitman et al., 1982). Thus, it may be ideal for social situations when an adult does not wish to disrupt ongoing social interactions between children. Ballard and Jenner (1981) attempted this with two elementary aged students who were demonstrating low levels of social behaviors during free play situations. The researchers targeted intervention during an unstructured time in the classroom when children had a choice of activities and were free to move around and talk as they pleased. During the “say” portion of the intervention, the children were asked the question, “What do we do to play with other children and be friends at school?” They were required to respond with the statements, “I go over to the other children,” “I do what they are doing,” and, “I smile.” The children were then able to play in the unstructured setting. Social interactions, defined as a reciprocal interchange between two children involving a directed social stimulus by one child and an observable response to that stimulus by the other children, were measured. Both children showed an increase in social interactions following the introduction of correspondence training.

Recess seems like another natural fit for a say-do intervention. Teachers or other school staff are often available to easily and naturally talk to a student before recess but would then like to avoid direct intervention during recess, both so as to not interrupt the natural flow of student interaction and to avoid the need for intensive staffing at recess. If a student could state his or her intention to interact with other children during recess before the event and a correspondence between his “saying” and his “doing” is then seen, interactions at recess could be increased without adult intrusion into ongoing recess social exchanges.

This study investigated the use of a say-do correspondence intervention with three first graders with ASD who had received high quality intensive social skills instruction in social skills groups three times a week for over a year at school, but who were still not demonstrating the social communication skills learned in the social skills groups during recess. In this study, the students were asked before recess to state who they would talk to at recess and then were reinforced afterwards if they had talked to the friend they had stated. Specifically we were interested in addressing the following research question: Does the use of correspondence training increase the frequency of social interactions at recess by elementary school children with ASD who have previously learned social behaviors during systematic social skills instruction in a social skills group with typically developing children?

**Method**

**Participants**

Participants were three children, two boys and a girl, with ASD who were participating in a
larger study of social and literacy skills (Kamps et al., 2014). All three children had received diagnoses of Autism Spectrum Disorder by independent medical professionals. Descriptive information about the three children can be seen in Table 1.

The three children attended different public elementary schools in the Pacific Northwest. As part of the larger research project, all three children had participated in a social skills group three times per week through kindergarten and were currently participating in a social group in first grade. These social groups, called Peer Networks social groups (Kamps et al., 2014), involved both the children with ASD and same-grade typically developing peers. For two of the children, Samuel and Jonathan, the groups involved three children on any given day: the target child and two typically developing peers. Different typically developing peers came on each of the three days of the week, resulting in six typically developing peers participating in Peer Network groups. The Peer Network social group for the third participant, Fran, involved four children on any given day: Fran, another student with ASD who also came three times per week, and two typically developing peers. As with the other groups, the typically developing peers came only once a week, resulting in six typically developing peers participating in Fran’s Peer Network.

The Peer Networks social groups were run using a manualized intervention focusing on verbal social communication skills, including asking questions, sharing, commenting on one’s play, commenting on a friend’s play, and play organizing. The importance of responding when another child initiated was stressed throughout the different skills. The social skills groups occurred three days a week for 30 minutes at a location and time selected by school staff. The groups were run by school staff and included special education teachers, instructional assistants, and speech language pathologists. The groups for two of the children involved in this study, Samuel and Fran, were run by special ed teachers; the group for the third child, Jonathan, was run by an instructional assistant. Training and on-going coaching for the intervention were provided by research personnel.

These three children were chosen for this study because, despite both this intensive social intervention involving typical peers for at least a year and a documented increase in verbal social communication during their social groups, none of the three target children demonstrated generalized use of the social skills they were learning when they were out at recess.

**Measures**

**Frequency of verbal exchanges.** The dependent variable in this study was the frequency of verbal exchanges between the participant and other children at recess. An exchange could be initiated by either the target child or another child. An exchange continued until 3 s had elapsed without any child in the interaction talking. Behavior was collected by observers on the playgrounds using IPads or iPhones equipped with the Behavior Tracker Pro application. The frequency of verbal exchanges was coded for 15 minutes during recess on every observation day.

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Age at time of participation</th>
<th>Adaptive Behavior (VABS-2)a</th>
<th>Language (CELF-Expressive Vocals: Age Equivalent)</th>
<th>SRS Tracher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samuel</td>
<td>M</td>
<td>6 years 8 months</td>
<td>72</td>
<td>8:11</td>
<td>66 (mild-to-moderate)</td>
</tr>
<tr>
<td>Jonathan</td>
<td>M</td>
<td>7 years 4 months</td>
<td>71</td>
<td>6:0</td>
<td>64 (mild-to-moderate)</td>
</tr>
<tr>
<td>Fran</td>
<td>F</td>
<td>7 years 5 months</td>
<td>67</td>
<td>7:2</td>
<td>81 (severe)</td>
</tr>
</tbody>
</table>

Note. aStandard scores (M = 100, SD = 15)
Reliability

Inter-observer reliability was obtained by having a second observer independently and simultaneously collecting frequency data on another iPad or iPhone, using the Behavior Tracker Pro software. Interobserver agreement was collected for 20%, 28.5%, and 14% of baseline sessions and 25%, 28.5%, and 36% of intervention sessions for Samuel, Jonathan, and Fran respectively. Interobserver agreement was calculated by dividing the number of agreements for each session by the total number of agreements and disagreements and then multiplying this number by 100. Interobserver agreement for Samuel averaged 96% (range, 88–100%), for Jonathan, 100%, and for Fran, 94% (range, 80–100%).

Setting

All phases took place in the public elementary schools attended by the participants at their regularly occurring recess time on the school playgrounds. Data was collected for the entire recess, which lasted for 15 minutes for each of children. The adult interactions before and after recess took place directly outside the door used for entering and exiting the playground. No data were collected during these interactions. No changes to the equipment, scheduling, or supervision/staffing of recess were implemented as part of this intervention.

Procedure

A combined concurrent and non-concurrent multiple baseline design (Harvey, May, & Kennedy, 2004) was used to assess the efficacy of the intervention. Intervention for the first two students took place in a concurrent multiple-baseline fashion during their first grade year. The third student, who entered the larger study a year later than the first two participants, was added non-concurrently a year later, also when she was in first grade. Sessions took place 1–2 times a week for each of the children.

Baseline. The child participated in recess in his or her usual manner. No special directions, schedules or visual aides were provided to the student. The researcher collected data on the participant’s frequency of social exchanges throughout the 15 minute recess. The researcher did not interact with the child or deliver any prompts. The researcher attempted to observe at a distance, making an effort to appear as if observing all children at play, not the focus child in particular.

Intervention. Immediately before the child went out to the playground, the researcher showed the child pictures of students who had participated in his or her social group and asked the child to name who he was going to talk to at recess. The child was allowed to either name a child or just point to a picture to identify who he wanted to talk to. The child was also allowed to name children not pictured (i.e., children not in his or her social group) although no child did. The child was then shown a collection of small, inexpensive toys and asked which one he would like to earn if he did what he said he would do. The researcher took this toy and kept it during recess.

The child then went out to recess to play. As in baseline, the researcher simply collected data on the frequency of social exchanges. Again no changes were made to the physical or social environment of the recess. The researcher also noted whether or not the child interacted with the student he had specified. As in baseline, the researcher did not interact with the child or deliver any prompts and again did her best to appear as if she was observing children in general, not the focus child in particular.

After recess the researcher debriefed with the child. The researcher asked the child if he had interacted with the child he had specified and the child received the reinforcer (the pre-selected toy and verbal praise) if he had. If he hadn’t, he was told, “To earn your toy, you need to talk to who you said you would talk to. You can try again tomorrow.”

Discontinuing pictures and tangible reinforcement. Once performance was stable, pictures and tangible reinforcement were discontinued. Before recess, the child was now simply asked to name a child that he would play with; no pictures were shown. During the debriefing at the end of recess, the child was praised if he talked to the person he had said he would talk to but no toy was given.
Results

As displayed in Figure 1, the frequency of social exchanges for the three students increased from baseline to intervention conditions.

Samuel (top panel) averaged 1.4 social exchanges during baseline. The frequency of his exchanges immediately increased to nine exchanges on the first day of the say-do intervention and remained variable but high for the rest of the intervention, averaging 15.4 exchanges over the next eight sessions. During the ninth session, pictures of peers to support his planning before the beginning of recess, and tangible reinforcement for implementing the plan that he made were discontinued. The correspondence intervention was continued, and verbal praise was provided for correct correspondence. Exchanges remained high, averaging 25.8 exchanges for the last four sessions.

Jonathan (middle panel) did not engage in any exchanges at all during baseline. His number of exchanges increased to three on the first day of intervention and climbed steadily to a peak of 15 exchanges before declining in the last two days of intervention before the end of the school year. Jonathan’s frequency of exchanges averaged 6.3 during intervention. Because of the end of school, the researcher was not able to implement the discontinuation of pictures and tangible reinforcement.

Fran (bottom panel) had a low, steady rate of social exchanges during baseline, averaging 4.7 exchanges during the 15 minute recess. Her frequency of exchanges immediately climbed to 11 on the first day of intervention and stayed above baseline rates for the next six sessions, averaging 12.2 exchanges per session. At this point, pictures and tangible reinforcement were discontinued. Exchanges stayed high, averaging 11.6 exchanges over the final four sessions.

Discussion

This study investigated whether a relatively simple intervention: a say-do correspondence where a child was asked who he would talk to at recess and then received reinforcement for doing so, would increase frequency of interactions at recess for children who had acquired these social communication skills in a social group but who were not generalizing the skills to recess. All three participants showed an increase in their social exchanges at recess as a result of the intervention. This confirms prior research indicating the effectiveness of the say-do correspondence procedure with children with autism (e.g., Morrison et al., 2002). The results also contribute to the growing literature showing positive interventions for recess settings for children with ASD (Lang et al., 2011).

This intervention seems ideal for a school recess environment. First, it fits well with the common staffing arrangements at school, where a teacher typically delivers her students to the playground before recess and greets them on their way in, but where staff on the playground is often minimal. It would be relatively simple for a teacher to interact briefly with a child before the child goes out to recess and then interact with him briefly again on the way in. Staff attendants on the playground need only observe whether the student has interacted with another specified student and briefly communicate this to the teacher at the end of recess. Even this level of recess time observation can presumably be discontinued once say-do correspondence is established. Second, this intervention also may be ideal in that it does not require prompting, intervening, or reinforcement while the social interactions between children are going on during recess, allowing the natural flow of child interactions to happen without interruption.

It is important to note, however, that in this study, we were not teaching the students new skills. Koegel, Vernon, Koegel, Koegel, and...
Figure 1. Frequency of social exchanges before and during intervention.
Paullin (2010) make the important distinction between skill and performance deficits and note that different approaches are needed for each. In this study, we were only trying to get the children to generalize skills that they had practiced intensively and for a long period of time (over a year in all cases) in their social groups; we were not trying to promote the acquisition of new skills. It seems unlikely that this low intensity intervention would have worked if the children did not have the skills of social interacting already in their repertoires.

It is also important to note that during the first phase of intervention, the focus child was shown pictures of students who were participating in his or her social groups when asked who her or she wanted to talk to, increasing the likelihood that these peers would be the focus of the child’s initial attempts at social exchanges. Thus, the focus child’s fledgling initiations at recess were to children who had received intensive practice at interacting with the child with ASD, potentially increasing the likelihood that the exchanges would be successful. It is unclear whether the same favorable outcomes would have been observed if the child with ASD had been encouraged to seek out just any child as their conversational target.

Limitations

Primary limitations of the study include the small sample size (only three participants), and the limited number of data points particularly for Jonathan. Monitoring of the intervention and maintenance over an extended period of time and for a larger number of participants would provide additional documentation of the intervention effects. Another potential limitation of this study is that it was conducted by researchers and we did not transfer implementation of the intervention over to school staff. Thus, we do not know for sure that the intervention would have been successful if school staff had been implementing it, nor do we know for sure that the school staff would agree on what we see as the feasibility of the intervention. Implementation of the intervention by researchers was appropriate for this study as we were establishing the effectiveness of the intervention in this setting. As our interest turns from evaluating effectiveness to evaluating questions of sustainability and ecological validity, however, it will be essential for school staff to implement the intervention that is being evaluated. Future research should involve using school staff to implement the intervention to answer these questions.

The presence of the researcher observing at recess may have been a contributing factor to the students’ performance. We do not know if the significant increases we saw in frequency of social exchanges would have continued if someone was not observing at recess and confirming correspondence with the child afterwards. Although tangible reinforcement was discontinued in this study, the researcher continued to observe the student at recess and gave the student verbal praise for talking to the targeted recipient. While the significant body of research on say-do correspondence suggests that the behavior increases would be sustained after the reinforcement for confirmed say-do correspondence ended, we did not establish this continuity in this study. Future research should establish correspondence methods similar to those used in this study and then fade out the observer at recess, confirming that increases are maintained in the absence of this confirmation of correspondence. This research should also investigate both if the intervention can be implemented intermittently and still be effective, and if the intervention continues to be effective when children are asked to make a plan every day, but teachers only debrief with them and provide feedback about their behavior intermittently.

To further increase the likely feasibility of this intervention in school environments, an interesting avenue for future research would be to assess whether adult time allocated to this intervention could be reduced even further through the use of self-management. Self-management has been shown to be an effective strategy for students on the autism spectrum (Lee, Simpson, & Shoran, 2007) and has been identified as an evidence-based practice for children and youth with ASD (National Professional Development Center on Autism Spectrum Disorders, n.d.). The say-do intervention investigated here required a brief interaction with an adult before and after re-
cess to first establish the goal and then confirm the “doing” of the stated goal. Perhaps once the intervention has been started and say-do correspondence has been established, a self-management system could be implemented, whereby the student shifts to using a computer (or other visual) to document before recess who they want to talk to (perhaps by choosing a picture or written name from an array or, if the student can write, typing a name) and then confirming via computer (or self-recording on a checklist) after recess whether or not he or she did what they had stated. Earned reinforcement could be delivered either immediately by the computer (perhaps via a fun animation) or the teacher could review the student’s performance later that day or even after several days and reinforce the student appropriately based on the correspondence between their initial choice and their self-reported actions at recess.

Use of the say-do correspondence procedure with children with ASD with more limited verbal skills might be another productive area of intervention research. Children in this study were highly verbal. Given that say-do correspondence has been effective using non-verbal “say” components (Whitman et al., 1982), the procedure might also be effective for lower functioning or less verbal children with ASD at recess.

**Conclusions**

This research has important implications for children with ASD and related disorders. The research is clear that using behavioral strategies, children with ASD can learn many skills across all developmental domains. The problem, however, is that these children often fail to demonstrate these newly acquired skills and behaviors in natural settings in which the skills would be most useful; that is, in the settings in which these skills and behaviors would have a true impact on the quality of life for these children with ASD and their families. Correspondence training is a relatively easy to use intervention that may be extremely effective in helping children with ASD learn to generalize newly acquired skills and behaviors across settings, activities, and people. As we learn more about how much correspondence training may be needed to promote and maintain generalized behavior, this school-friendly intervention may become a staple in the intervention history for every child with ASD or related disorders that have challenges in the area of generalization.

In summary, the data reported here suggest that correspondence training is an effective and sustainable strategy that can be used to enhance the social behavior of young children with ASD. Correspondence training may be a simple strategy to help students learn to demonstrate their newly acquired skills across settings. This type of instruction may be especially important for children with ASD, who have well documented difficulties with generalization, and may be an effective strategy to help classroom teachers enhance their students’ abilities across settings and staff.

**References**


An Investigation of the Effects of CRA Instruction and Students with Autism Spectrum Disorder

Shaunita Stroizer
Valdosta State University

Vanessa Hinton, Margaret Flores, and LaTonya Terry
Auburn University

Abstract: Students with Autism Spectrum Disorders (ASD) have unique educational needs. The concrete representational abstract (CRA) instructional sequence has been shown effective in teaching students with mathematical difficulties. The purpose of this study was to examine the effects of the CRA sequence in teaching students with ASD. A multiple baseline across behavior design was used in assessing the effects of CRA to three elementary students with ASD over four weeks of instruction. A functional relation was demonstrated between CRA and three behaviors: addition with regrouping, subtraction with regrouping, and the multiplication facts zero through five. The results and implications are discussed further.

The number of children eligible for special education services under the category of Autism Spectrum Disorder (ASD) has increased (Centers for Disease Control and Prevention, 2012). According to Rice (2009) children identified as having ASD increased 60% among males and 48% among females from the year 2002 to 2006. No Child Left Behind (2002) and the Individuals with Disabilities Educational Improvement Act (2004) make it clear that educators must provide research based instruction for students with disabilities, and that students with disabilities must show proficiency in the general education grade level standards. This includes students with ASD. The National Council of Teachers of Mathematics (NCTM, 2000), the National Mathematics Advisory Panel (NMAP, 2008), and the Common Core Standards for Mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) call for in-depth instruction of whole numbers as critical for elementary aged students. In particular, the NCTM, NMAP, and Common Core Standards for Mathematics suggested elementary students must develop an understanding of the meaning of the basic operations of addition, subtraction, multiplication, and division as well as build fluency with the standard algorithms for addition, subtraction, multiplication, and division. The NCTM (2006) set focal points for mathematic instruction for prekindergarten to Grade 8 which included understanding numeric operations and their relationship to each other. The NMAP also explained that practice with conceptual understanding of whole numbers allows students to achieve automaticity of basic skills which is the fast, accurate, and effortless processing of content information. This automaticity frees up working memory for more complex aspects of problem solving. Conceptual understanding of numbers and operations demonstrated through representation of objects is an integral component of the Common Core Standards for Mathematics. Since students with disabilities, including students with ASD, are expected to make progress within the general education curriculum through the Common Core Standards for Mathematics, it is important that effective supplemental instructional methods are explored.

To date, researchers have demonstrated that a sequence of instruction referred to as concrete-representational-abstract (CRA) as
an effective means to teach and supplement mathematics instruction for students with disabilities. The CRA sequence involves three levels of instruction. The first level, the concrete, uses manipulatives to promote conceptual understanding; problems are completed through manipulation of objects. The second level, the representational level, involves completion of problems using drawing to represent numbers instead of manipulating objects. This level of instruction serves as a bridge to the final level, the abstract level in which a student uses numbers to solve problems. Instruction within each level is explicit, meaning that it includes an advance organizer, teacher demonstration, guided practice, independent practice, immediate feedback, and a post organizer.

Skills in which CRA instruction was demonstrated as a successful intervention for students with disabilities include integers, fractions, and algebra concepts (Butler, Miller, Crehan, Babbitt, & Pierce, 2003; Maccini & Hughes, 2000; Maccini & Ruhl, 2000; Witzel, Mercer, & Miller, 2003). Additionally, CRA has been shown effective in teaching basic early mathematic computation for students who struggle with learning mathematics (Flores, 2009; Harris, Miller, & Mercer, 1995; Mercer & Miller, 1992; Morin & Miller, 1998). CRA may possibly be an effective method of teaching basic computation to students with ASD (Donaldson & Zager, 2010) because research indicates that they prefer visual modes of communication and have particular difficulty understanding abstract concepts (Boutot & Myles, 2011; Ganz, Earles-Vollrath, Cook, 2011). Due to the current educational requirements (IDEA, 2004; NCLB, 2002) many students with ASD are accountable for acquiring general education standards in which knowledge of basic mathematics computational skills are key components. It is possible CRA could be a means for students with ASD to achieve general education curriculum standards.

Kroesbergen and Van Luit (2003) examined the effectiveness of mathematical interventions for students with disabilities, specifically interventions that involved explicit instruction, cognitive self-instruction, and assisted instruction. With regard to basic math computation, they found that interventions that utilize explicit instruction were more effective than cognitive self-instruction and assisted instruction for students with disabilities. Students with ASD could benefit from explicit instruction for the following reasons: (a) explicit teaching is teacher directed instruction that is organized and task oriented, (b) concepts are presented in a clear and direct manner, and (c) students respond to instruction and receive immediate feedback (Miller 2009; Peterson, Mercer, and O’Shea, 1988; Witzel et al., 2003).

Concrete Representational Abstract Instruction

Peterson et al (1988) combined explicit instruction and the use of CRA to teach place value to the tens place. Students who participated in the study received special education services for specific learning disability (SLD). There were 20 males and 4 females ranging in age from 8 to 13. A multiple baseline across subjects design was used to investigate explicit CRA instruction and student achievement regarding place value. Probes and maintenance data revealed that students mastered place value identification with an 80 percent or greater level of accuracy.

Mercer and Miller (1992) built upon Peterson et al.’s work and developed a curriculum called Strategic Math Series (SMS) utilizing explicit instruction and the CRA sequence to teach place value and basic math facts to students with disabilities. Mercer and Miller field tested SMS with 109 elementary students of whom 102 had a learning disability, five students had an emotional behavior disorder, and two were at risk. Total mean scores demonstrated that the average gain across skills was 59%.

Harris et al (1995) evaluated the effectiveness of teaching multiplication facts using explicit teaching and CRA instruction to elementary students with disabilities in the general education classrooms. Participants of this study included 13 second graders with disabilities and 99 second grade students without disabilities. The performance of students with disabilities improved; it was equivalent or slightly improved compared to that of their normally achieving peers during the phase of instruction that required demonstration of conceptual understanding of the multiplication process. With a strategy for counting ob-
jects within groups, subjects with disabilities accurately completed basic multiplication problems during independent computation and problem solving practice activities.

Morin and Miller (1998) extended the study of Harris et al. using multiplication and the SMS curriculum. The purpose of this study was to evaluate the effectiveness of teaching multiplication facts and related word problems using the CRA teaching sequence to middle school students with intellectual disabilities. A multiple baseline across subjects design was used to demonstrate experimental control and a functional relation between interventions that used the CRA sequence, and multiplication facts the students learned to solve word problems. All student improved their performance to 90% accuracy. The findings of this investigation suggest that students with intellectual disabilities were able to learn to solve multiplication facts and related word problems using the CRA teaching sequence and explicit instructional procedures.

Flores (2009) expanded on Miller and Mercer’s research using explicit instruction and CRA to teach students how to subtract with regrouping. Participants were six third-grade students who were all failing mathematics. The study used a multiple probe design to evaluate the efficacy of CRA instruction for teaching subtraction with regrouping. A functional relation was demonstrated between CRA and subtraction skills. All five students met the criterion of writing 20 digits on three consecutive 2 minute curriculum-based measures. Of the five students, four maintained their performance 4 weeks after the end of instruction.

Similar to Flores (2009), Kaffar and Miller (2011) investigated the effects of CRA instruction but included the mnemonic titled “RENAME” for subtraction with regrouping. Participants who received the instruction utilizing CRA consisted of eight students with math difficulties and three students with disabilities. Twelve students were in the control group and received instruction using a basal program. Both groups made gains in subtraction with regrouping, however, the gains made by the students who received instruction with the CRA sequence was greater. Students’ mean percentage computation scores increased from 49% to 90% for the treatment group while the comparison group’s scores increased from 66% to 72%.

All of the CRA literature reviewed included explicit instruction (an advance organizer, teacher modeling, teacher guidance, independent practice, and a post organizer) and demonstrated the CRA sequence to be effective in teaching the participants mathematics operations (Harris et al., 1995; Flores, 2009; Kaffar & Miller, 2011; Mercer & Miller, 1992; Morin & Miller, 1998; Peterson et al., 1988). Currently, studies have not been implemented to examine the effects of instruction utilizing an explicit instruction curriculum with the CRA sequence for students with ASD. This conclusion was drawn after searches that included key words: “mathematics,” “mathematics interventions,” concrete-representational-abstract,” “CRA,” “autism spectrum disorder,” “autism,” “Asperger Syndrome,” “mathematics instruction,” and “explicit mathematics instruction,” using the following databases: Psych Info, ERIC and Education Research Complete.

Mathematics instruction for students with ASD might be particularly beneficial for the following reasons. Research has shown that students with ASD prefer instruction through a visual modality (Boutot & Myles, 2011). The use of manipulative objects and pictures provides students with a visual and concrete representation of the abstract concepts associated with numerical operations. Students with ASD perform better when instruction is structured and predictable (National Autism Center, 2009). The general education curriculum through the Common Core Standards for Mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) requires students to represent numbers and operations. Instruction using the CRA sequence provides students with explicit instruction, implemented in a structured manner that guides students through each component of a complex process of representing operations. Students with ASD represent a growing segment of learners who participate in the general education curriculum and there is currently no research regarding supplemental interventions or instructional strategies to inform classroom practice.

Therefore, the purpose of this study is to examine the effects of the CRA sequence on
the mathematics performance of students with autism spectrum disorders. The following questions guided the study. What are the effects of CRA instruction on students with ASD’s addition with regrouping performance? What are the effects of CRA instruction on students with ASD’s subtraction with regrouping performance? What are the effects of CRA instruction on students with ASD’s multiplication performance?

Method

Participants

Three male elementary students participated in the study. Larry was entering Grade 3 and Devin and Marvin were entering Grade 4 in the upcoming school year. The participants’ characteristics are included in Table 1.

The criteria for participation were: (a) eligibility for special education services under the ASD category; (b) participants received special education services in the area of mathematics; (c) participants demonstrated difficulty with mathematics achievement as assessed using the KeyMath III (Connolly, 2007); and (d) parent permission to participate in the study. Students were recruited through letters sent to all parents with their application packets to the program. Of the students for whom permission was obtained, assessments were administered to determine eligibility for participation. In addition, cognitive ability scores were obtained using the Kaufman Brief Intelligence Test 2nd edition (K-BIT, Kaufman & Kaufman, 2004).

Larry received instruction in the general education classroom for the majority of the school day during the regular school year. Larry qualified for special education services under the category of ASD according to state and federal guidelines and was originally diagnosed at age three by a pediatric neurologist. He performed significantly below average in mathematics achievement according to the KeyMath III (Connolly, 2007). Devin received instruction in the general education classroom for the majority of the school day during the regular school year. Devin qualified for special education services under category of ASD according to state and federal guidelines and was originally diagnosed at the age of two years and four months by a team of professionals at a medical school center for developmental and learning disorders. He performed significantly below average range in mathematics achievement according to the KeyMath III. Marvin received instruction in the general education classroom for the majority of the school day during the regular school year. Marvin qualified for special education services under category of ASD according to state and federal guidelines and details regarding his original diagnosis were unavailable. He performed significantly below the average range in mathematics achievement according to the KeyMath III.

Setting

The study took place within a university-sponsored extended school year program in a rural region of the Southeastern United States. The program was one month in length and students attended the program five days a week for three hours. Instruction was provided in reading, written expression, and mathematics

---

### Table 1: Participant Characteristics

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
<th>Age</th>
<th>Cultural Background</th>
<th>Cognitive Compositea</th>
<th>Math Achievement and Grade Equivalentb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larry</td>
<td>male</td>
<td>10</td>
<td>White</td>
<td>65</td>
<td>65/1.4 Grade level</td>
</tr>
<tr>
<td>Devin</td>
<td>male</td>
<td>10</td>
<td>White</td>
<td>73</td>
<td>73/1.6 Grade level</td>
</tr>
<tr>
<td>Marvin</td>
<td>male</td>
<td>8</td>
<td>Latino</td>
<td>55</td>
<td>68/1.5 Grade level</td>
</tr>
</tbody>
</table>

a Standard score from Kaufman Brief Intelligence Test (Kaufman & Kaufman, 2004).
b Standard score from KeyMath III (Connolly, 2007).
according to the students’ goals within their individualized educational programs (IEP). Mathematics instruction was provided daily for 60 minutes by a certified special education teacher who was not a researcher. The teacher had a Master’s degree and three years teaching experience in an elementary school with students with ASD. The teacher was chosen because she was knowledgeable in the intervention. Prior to beginning, the teacher demonstrated a lesson implemented with 100% accuracy for each CRA level and skill.

Materials

The research materials included instructional items such as learning sheets, manipulative items and probes for assessment purposes. The progress monitoring probes were sheets of paper with nine problems. The probes were created through the use of a website named Intervention Central in which curriculum based assessments to monitor student progress were generated. There were four different probes that were alternated for each different skill area: four addition probes (two digit by two digit) that required regrouping, four subtraction probes (two digit by two digit) that required regrouping, and four single digit multiplication probes that included the digits 0–5. For example, one sheet consisted of computational problems for addition with regrouping, one sheet consisted of computational problems for subtraction with regrouping, and one sheet consisted of computational problems for multiplication problems that include the digits 0–5. Participants completed the probes during independent morning work that lasted between five and ten minutes before daily instruction began. Once all participants completed the three probes for their morning work, instruction began.

Measures to ensure content validity for the probes were conducted for each skill area. Pools of the computation questions for each skill area were distributed to five teachers to review. All five teachers had at least a master’s degree from an accredited university with an average of 5 years of experience in teaching mathematics to elementary or middle school students. The reviewers were asked to score each problem for each skill area according to its relevance to the content using a 4-point scale to avoid having neutral and ambivalent midpoint. The item relevance continuum used was 1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, and 4 = highly relevant. Then for each item, the item-level content validity index (I-CVI) was computed as the number of experts giving a rating of 3 or 4 (thus dichotomizing the ordinal scale into relevant and not relevant), divided by the total number of experts (Lynn, 1986). The I-CVI for the pool items was calculated as 1.00 for addition with regrouping, 1.00 for subtraction with regrouping, and 1.00 for multiplication with digits 0–5.

To further examine the items in each probe, teachers were asked to rate the same problems as the following: not difficult, somewhat difficult, quite difficult, or highly difficult for the average elementary school student who was learning the computation problem. The difficulty level of the questions required consistency of the majority of the raters on each problem. Based on the survey, 21 of the 27 problems for addition with regrouping were rated as quite difficult and six problems were rated as somewhat difficult. For subtraction with regrouping, all 27 problems were rated as quite difficult and for multiplication with numbers zero through five, 10 problems were rated as quite difficult and 17 problems were rated as somewhat difficult.

Reliability of the probes was also examined for each skill area. The researchers created an assessment for each skill area (one assessment for subtraction with regrouping, one assessment for addition with regrouping, and one assessment for single digit multiplication computation with numbers 0–5). Each assessment contained 75% or more of the problems from the probes created for the participants. The researchers administered the three assessments to 20 college level students from a major four-year university. The assessments were untimed and 16 of the assessments were completed (four students chose not to volunteer). Results from the internal consistency test revealed Cronbach’s Alpha Coefficient of $r = .72$ for addition with regrouping, $r = 1.00$ for subtraction with regrouping, and $r = .72$ for basic multiplication computation with digits 0–5.

The CRA intervention materials included base-ten blocks made of foam and learning sheets in which problems were divided into three instructional sections: (a) model, (b)
guided practice, and (c) independent practice. Additionally, the written computation of addition, subtraction, and multiplication sub-tests of the Operations domain of the KeyMath III (Connolly, 2007) and Kaufman Brief Intelligence Test II (KBIT-2; Kaufman & Kaufman, 2004) were administered to assess mathematics achievement and students’ cognitive functioning. A review of the content of the KeyMath III by the NCTM publications concluded the instrument reflected the content and process standards described in the NCTM Principles and Standards for School Mathematics (2000). Additionally, the KeyMath III scores were correlated with scores on the KeyMath Revised, Normative Update: A Diagnostic Inventory of Essential Mathematics; the Kaufman Test of Educational Achievement, Second Edition; Iowa Tests of Basic Skills instrument; the Measures of Academic Progress; and the Group Mathematics Assessment and Diagnostic Evaluation. An intercorrelation was reported for the total test and Operations domain as \( r = .92 \) for individuals Grades 3 through 5. Also, the internal consistency coefficient for the operation domain was reported \( r = .89 \) (Form A) and \( r = .90 \) (Form B) for individuals in Kindergarten through Grade 5. The KBIT-2 was used to gain a quick estimate of intelligence. It assesses verbal and nonverbal intelligence in people from 4 to 90 years of age. The KBIT-2 provides scores for verbal intelligence, nonverbal intelligence, and a composite of overall intelligence. Correlations of the KBIT-2 to the Wechsler Abbreviated Scale of Intelligence-III (WASI-III) and the Wechsler Abbreviated Scale of Intelligence-IV (WASI-IV) are strong at .76 and .77 respectively. Comparisons showed Full Scale and Performance scores about 4.5 points and 7 points higher on the Wechsler Abbreviated Scale of Intelligence than the corresponding KBIT-2 scales. The KBIT-2 composite of overall intelligence has an internal consistency coefficient of .93 across all ages.

**Design and Procedure**

A multiple baseline across three separate behaviors was utilized to evaluate the effects of CRA for students with ASD. Multiple baseline was used to ensure enough data were collected for the study because the program was only four weeks in length. Data were collected across the behaviors of addition with regrouping, subtraction with regrouping and multiplication facts zero to five through probes. Students completed the probes during morning work before mathematics instruction began. Probes were administered daily for each mathematical skill collecting baseline and intervention data. Probes were scored based on correct answers. During the baseline phases, no instruction was provided regarding the target mathematical skill; therefore this design was used to observe student performance with and without CRA intervention.

Upon introduction of the CRA interventions, mathematics instruction was provided on a daily basis for 5 days a week. Instruction was provided in a small group and implemented by the lead teacher of the classroom. Instructional sessions lasted approximately 20 minutes in the beginning but increased to 60 minutes (20 min per behavior) as participants reached criterion and instruction was provided for subtraction with regrouping, then single digit multiplication. All of the students began instruction together, but when one student reached criterion for the first behavior, he received instruction for the next behavior for an additional twenty minutes. Since students progressed differently, instructional groups fluctuated in size from three to one.

The criterion for phase change was six of nine problems correct because the goal of the study was to provide intensive instruction in a brief amount of time. The researchers chose six because it indicated progress, at least a fifty percent increase in performance over baseline. Instruction continued after the criterion for phase change was reached. For example the student continued instruction in addition with regrouping after instruction in subtraction with regrouping began. Additionally, the mathematics skills of addition with regrouping and subtraction with regrouping are complementary skills that can be taught at the same time. Instruction in basic multiplication facts does not require regrouping as prerequisite and can be taught at the same time. The purpose of the summer program was remediation; therefore there was an emphasis on efficient and effective instruction. There were a total of 20 days in which instruction took place. The first week involved assessment, placement into small groups for instruction,
and baseline data collection. Data were taken for all mathematic skills (i.e., addition and subtraction with regrouping, and basic multiplication). Once three stable data points were established, instruction began for addition with regrouping. Once a student met the criterion of six problems correct, addition instruction continued; however, the next phase of the study began in which the intervention for subtraction with regrouping was included. Once a student met the criterion of six problems correct for subtraction with regrouping, the third phase of the study began in which a student received instruction for multiplication facts zero through five. Students progressed differently; therefore instructional groups fluctuated in size from three to one.

Instructional Procedures for Addition and Subtraction with Regrouping

CRA instruction for addition and subtraction was modeled after the SMS addition and subtraction with regrouping curriculum used in previous research (Flores, 2009; Kaffar & Miller, 2011; Mercer & Miller, 1992). The concrete phase of instruction included the first three lessons. For each lesson, a sample script and learning sheets guided the teacher through the instructional sequence. During these lessons, the teacher modeled problem solving by manipulating concrete objects (i.e., base-ten blocks) to solve problems. The addend or minuend was represented with the base-ten blocks and beginning in the ones place, the other addend or subtrahend was combined or taken away, trading ten ones for one ten or one ten or ten ones when regrouping was necessary. This continued for the tens and hundreds places. Then, the teacher guided students through turn-taking and together they solved problems. This was followed by independent practice in which the students completed problems without assistance. The lesson ended with a post organizer in which the lesson was reviewed.

Following the SMS curriculum procedures, representational level instruction did not begin until students completed independent practice problems with 80% accuracy. Representation level lessons proceeded using similar methods, except the problems were represented using drawings (tallies for ones, long lines for tens, and squares for hundreds). Lesson seven introduced the “RENAME” strategy (Kaffar & Miller, 2011). “RENAME” included six steps. This involved (a) read the problem, (b) examine the ones, (c) note the ones, (d) address the tens column, (e) mark the tens column, and (f) examine and note the hundreds and exit with a quick check. Abstract instruction started in lesson eight and continued until the end of the program in which RENAME helped students remember problem solving steps using numbers only.

Instructional Procedures for Multiplication Facts 0–5

The procedures for CRA instruction for multiplication were those associated with the SMS Multiplication Facts curriculum. They were similar to those used to teach addition and subtraction with regrouping except students used manipulative items and pictures to represent repeated addition. For example, the problem 2x2, was represented as two groups of two objects. In addition, once students mastered operations using objects and pictures, they learned “DRAW” instead of “RENAME”. The “DRAW” strategy has four steps: (a) discover the sign, (b) read the problem, (c) answer with a conceptual representation, and (d) write the answer.

Integrity and Inter-Observer Agreement

Data for integrity were collected throughout each of the instructional conditions within the study to ensure that research and intervention procedures were implemented correctly. Integrity checklists were completed in which two researchers observed administration of probes and instruction as it was taking place as well as digital recordings of the probe administration and instructional lessons. Integrity checks were conducted during 27% of the sessions pertaining to baseline and addition with regrouping, 30% for the sessions pertaining to baseline and subtraction with regrouping, and 60% for the sessions pertaining to baseline and multiplication facts zero through five. The checklist used to measure integrity included behaviors such as: the instructor provides students with a blank probe sheet, instructs students to complete as many problems
as possible, and does not give answers; the instructor says to students what they will be doing and why, the instructor demonstrates how to solve the problems using objects, pictures, or numbers; students and the instructor solve problems together; the instructor tells students to solve problems independently and does not offer answers; the instructor provides feedback regarding student responses; and lastly, the instructor closes with a positive statement about student’s performance in the feedback process and mentions future expectations. To calculate integrity, the researcher took the total number of agreements between two researchers and divided the number by the total number of observations; then multiplied by 100 for each condition (Poling, Method, & LeSage, 1995). The level of procedural integrity for instructional sessions was calculated as 88% for addition with regrouping condition, 97% for subtraction with regrouping condition, and 81% for multiplication facts zero through five condition. Procedural integrity was calculated at 81% for the multiplication condition because the instructor did not explain why the students were doing multiplication and did not explain future expectations consistently during multiplication lessons.

Inter-Observer Agreement was conducted for 100% of the probes administered as well as the integrity checklists. Probes were collected before each instructional lesson. Probes were scored by the instructor and later scored again by a researcher. To calculate inter-observer agreement, the total number of agreements between a researcher and the instructor was divided by the total number of agreements plus disagreements; then multiplied by the number 100. Inter-observer agreement was 100% for probes and 98% for the integrity checklists.

Social Validity

Social validity was addressed through a closed and open ended questionnaire after the study. The teacher who implemented the instruction answered questions regarding the efficacy of the intervention and recommendations for the intervention. The teacher’s feedback indicated that the SMS curriculum taught students with ASD to complete the computational problems and improved their skills. In addition, the teacher indicated the curriculum was worth the time and instructional effort, would use the intervention again, and recommended its use to others. Lastly, the teacher wrote she liked the universal approach of the curriculum when presenting math skills with concrete, representational, and abstract strategies.

Results

A multiple baseline across behaviors design was utilized to evaluate the effects of CRA for students with ASD. Data were interpreted by visual inspection and the following were noted: overlap between baseline and treatment, slope of each treatment data path, and number of data points from the beginning of treatment to criterion. Results for Larry, Devin, and Marvin are summarized in Figures 1, 2, and 3.

Larry’s baseline data were stable across all behaviors, zero correct problems on all probes. Devin’s baseline performance on addition with regrouping was more variable ranging from zero problems correct to two problems correct, however his performances on the subtraction with regrouping and multiplication probes were stable at zero problems correct. Marvin’s performances on addition and subtraction with regrouping probes were stable with zero problems correct on each, yet his performance on the multiplication probes had some variation ranging from zero problems to one problem correct.

Performance after Implementation

Larry. Larry received fifteen lessons for addition with regrouping before the Summer Program ended and completed a total of seventeen probes. He reached criterion for addition with regrouping after two probes. There was an immediate change in performance between baseline and CRA instruction no overlapping data points across the baseline and instructional phases. The instructional phase data points show an upward trend which indicates steady improvement. Larry received twelve lessons for subtraction with regrouping and completed a total of eighteen probes. He reached criterion after three probes. There
was a change in performance level; however the first data point overlapped with baseline. The instructional phase data points indicate an upward trend. Larry received eight lessons for multiplication facts and completed a total of nineteen probes. He reached criterion after three probes. There was a change in performance between baseline and CRA, no overlapping data points, and an upward trend.

Devin. Devin received thirteen lessons for addition with regrouping before the Summer Program ended and completed a total of 16 probes. He reached criterion for addition with regrouping after two probes. There was an immediate change in performance between baseline and CRA instruction, no overlapping data points, and an upward trend which indicates steady improvement.

Marvin. Marvin received 13 lessons for addition with regrouping before the Summer Program ended and completed a total of 15
probes. He reached criterion for addition with regrouping after one probe. There was an immediate change in performance between baseline and CRA instruction and no overlapping data points between the baseline and instructional phases. The instructional phase data points show an upward trend which indicates steady improvement. Marvin received 11 lessons for subtraction with regrouping and completed a total of 15 probes. He reached criterion after four probes. There was a change in performance level; however the first data point overlapped with baseline performance. The instructional phase data points show an upward trend. Marvin received eight lessons for multiplication facts zero to five and completed a total of 18 probes. He reached criterion after two probes. There was an immediate change in performance between baseline and CRA instruction, no overlapping data points, and upward trend which indicates steady improvement.

**Effect size**

Tau-U was calculated for each student; this form of analysis combined non-overlapping data points between phases with trend within the intervention phases while accounting for any trend within baseline (Parker, Vannest, Davis, & Sauber, 2011). For Larry, there were no significant trends within baseline phases. In comparing baseline and intervention phases for addition with regrouping, a strong effect was indicated (Tau-U = 1). In comparing Larry’s baseline and intervention data for subtraction with regrouping, a strong effect was indi-
The comparison of Larry’s baseline and intervention phases for multiplication indicated a strong effect (Tau-U = 1). Overall, the intervention had a strong effect across all phases (Tau-U = .97).

For Devin, there were no significant trends within baseline phases. In comparing baseline and intervention phases for addition with regrouping, a strong effect was indicated (Tau-U = 1). In comparing Devin’s baseline and intervention data for subtraction with regrouping, a strong effect was indicated (Tau-U = .90). The comparison of Devin’s baseline and intervention phases for multiplication indicated a strong effect (Tau-U = 1). Overall, the intervention had a strong effect across all phases (Tau-U = .97).

For Marvin, there were no significant trends within baseline phases. In comparing baseline and intervention phases for addition with regrouping, a strong effect was indicated (Tau-U = 1). In comparing Marvin’s baseline and intervention data for subtraction with regrouping, a strong effect was indicated (Tau-U = .90). The comparison of Marvin’s baseline and intervention phases for multiplication indicated a strong effect (Tau-U = 1). Overall, the intervention had a strong effect across all phases (Tau-U = .97).

**Discussion**

In the current study, the authors investigated the effectiveness of CRA for teaching addition and subtraction with regrouping and multiplication facts zero to five to students with ASD.
A functional relation was demonstrated between the CRA instruction and the behaviors of addition and subtraction with regrouping and multiplication facts zero to five with three participants. The students participated in each level of instruction without modifications, and moved through the phases of instruction with relative ease. All three students demonstrated steady progress across the three skill areas. The students’ positive outcomes may have been related to the visual nature of instruction. This involved watching the teacher move objects to demonstrate the mathematical concept of addition then allowing the students to manipulate objects themselves demonstrating their understanding. For example, Marvin improved after one day of instruction for addition with regrouping. Marvin demonstrated what addition with regrouping looked like with objects and received immediate feedback. Therefore, when he solved problems using numbers only, his previous demonstrations of addition with regrouping at the conceptual level using manipulative items and pictures may have made the task easier. Perhaps Marvin’s sudden increase in performance was related to the correction an error pattern or conceptual misunderstanding, but CRA instruction corrected his misconceptions or lack of procedural understanding. This might also explain the sudden changes observed for Devin in the multiplication condition. Based on background information and their individualized education programs, the students had received instruction in these areas during their regular school year using general education curriculum materials. Perhaps instructional design and the use of the CRA sequence were more conducive to learning and consistent with research related to students with ASD’s preference for visual input (Ganz et al., 2011). Also, Instruction was explicit and followed predictable patterns of demonstration, guided practice, and immediate feedback, consistent with instructional standards for students with ASD (National Autism Center, 2009).

Consistent with previous research (Harris et al., 1995; Flores, 2009; Kaffar & Miller, 2011; Mercer & Miller, 1992) the effects of the CRA instruction resulted in mathematical computation gains. The students’ pretests indicate poor performance regarding addition and subtraction with regrouping as well as basic multiplication. The CRA instructional sequence provided students with clear expectations and the scaffolding regarding the conceptual meaning of operations to support the procedural knowledge needed to complete the problems. Perhaps the students made gains because instruction involved focus on conceptual knowledge through manipulation of objects and immediate feedback from the teacher. This is consistent with Donaldson and Zager’s (2010) argument that CRA’s explicit nature and scaffolding from concrete to abstract may benefit students with ASD. The CRA sequence provided explicit instruction, appealing to students’ visual preferences (Boutot & Myles, 2011; Ganz et al, 2011).

The skills gained by the students in this study are consistent with learning standards within the Common Core Standards for Mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). The students demonstrated increased conceptual knowledge of operations, showing knowledge of place value and its relation to operations through regrouping. In addition, students demonstrated knowledge of multiplication and its relation to addition by repeatedly adding groups. This is significant because these skills will allow for progress within the general education curriculum with more complex concepts.

This study extends the current literature (Flores 2009; Harris et al., 1995; Kaffar & Miller, 2011; Morin & Miller, 1998) by investigating the effects of the CRA instructional sequence and mathematics achievement for students with ASD. This is significant in that it provides an initial investigation of interventions for students with ASD. With the introduction of the Common Core Standards for Mathematics (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), it is important to continue to conduct research regarding effective interventions and to also include students with disabilities who will participate in general education instruction such as students with ASD.

Limitations and Future Research

The research design presents a limitation to the present study because the CRA instructional sequence was not compared with an-
other mathematics program. Therefore, there may be other instructional programs or interventions as or more effective. Another limitation of this study is that social validity data were not collected directly from the students who participated in the study; only the teacher provided social validity data. In future studies, data from the students must be collected and reported. Lastly, generalizability of the results across settings is a limitation because the setting is outside the general education classroom and involves small group instruction. It is unclear if the same effects would be replicated in a larger group, inclusive setting. Also, due to the limited time frame of the program, maintenance data were not collected. Therefore, the long term effects of the instructional intervention are unclear.

Future research should be conducted in order to extend and replicate findings related to CRA and increased computation performance for students with ASD. Since there is little research for this population related to mathematics interventions, it is important to investigate the efficacy and efficiency of mathematics interventions such as CRA. Although CRA has been shown to be effective for students with high incidence disabilities, there may be particular aspects of CRA or other interventions that are effective for students with ASD who participate in the general education curriculum. Another area of future research may be an examination of the long term effects of CRA instruction for students with ASD. Researchers should investigate the effects of CRA in an inclusive setting for students with ASD and with larger groups of participants. In addition, research should be conducted to compare the effects of CRA with other mathematical programs that provide instruction for students with ASD.

References
students with learning problems in math to acquire, understand, and apply basic math facts. Remedial and Special Education, 13, 19–35.
Received: 19 December 2013
Initial Acceptance: 22 February 2014
Final Acceptance: 22 April 2014
Effect of Instruction with the Self-Determined Learning Model of Instruction on Students with Disabilities: A Meta-Analysis

Suk-Hyang Lee
Ewha Womans University

Michael L. Wehmeyer and Karrie A. Shogren
University of Kansas

Abstract: The Self-Determined Learning Model of Instruction (SDLMI) has been identified as an evidence-based practice to support educators to teach students with disabilities to engage in self-regulated learning leading to enhanced self-determination, attainment of academic and functional goals, and enhanced access to the general education curriculum. This article reports the results of a meta-analysis of the efficacy of applying the SDLMI as an intervention to enhance access to the general education curriculum or transition-related outcomes. Fifteen single-subject research studies examining the efficacy of the SDLMI as an intervention for students with disabilities were synthesized and the impact was analyzed across intervention and participants’ characteristics using the percentage of non-overlapping data (PND) metric. The results of this analysis add to the evidence for the efficacy of implementing the SDLMI as an intervention to promote academic and functional goal attainment for students with disabilities. Implications and directions for future research are discussed.

Promoting self-determination has received attention from practitioners and researchers in special education and related disciplines as a means to enhance access to the general education curriculum for students with disabilities as well as to promote more positive transition-related outcomes for these youth (Agran et al., 2005; Lee, Wehmeyer, Palmer, Soukup, & Little, 2008; Palmer, Wehmeyer, Gipson, & Agran, 2004; Ward & Kohler, 1996; Wehmeyer, 2007; Wehmeyer, Field, Doren, Jones, & Mason, 2004). Although there have been significant progress in recognition of the importance of self-determination to more positive outcomes for students with disabilities among practitioners, there remains a need to translate knowledge-to-practice about methods, materials, and strategies educators can use to promote the self-determination of students with disabilities (Agran, Snow, & Swaner, 1999; Lee, 2009a; Wehmeyer, Agran, & Hughes, 2000). Much of the research on teacher knowledge about and awareness of self-determination has been in the U.S., but research in Korea, where efforts to promote self-determination have been given considerable attention (Bae & Wehmeyer, 2003; S. H. Lee & Wehmeyer, 2004, 2007; Y. Lee & Wehmeyer, 2008; Seo et al., 2012) suggests a similar situation. Lee (2009b) investigated perceptions of and issues pertaining to the self-determination of students with disabilities and instructional practices related to self-determination through in-depth interviews with 17 special education teachers and 10 parents of students with disabilities. The results supported an ongoing need to provide teachers and parents with opportunities to learn about the concept of self-determination and how to promote the self-determination of students with disabilities.

That promoting self-determination is important to students with disabilities is now a matter of evidence and not just speculation. Wehmeyer and colleagues (Shogren, Wehmeyer, Palmer, Riftenbark, & Little, 2015; Wehmeyer, Palmer, Shogren, Williams-Diehm, & Soukup, 2013)
conducted a randomized trial control group study of the effect of interventions to promote self-determination on the self-determination of high school students receiving special education services under the categorical areas of intellectual disability and learning disabilities. Students in the treatment group (n = 235) received instruction using a variety of instructional methods to promote self-determination and student involvement in educational planning meetings over three years—which will be detailed in a subsequent section—while students in the control group (n = 132) received no such intervention. The self-determination of each student was measured using two norm referenced measures of self-determination across three measurement intervals (At Baseline, After 2 Years of Intervention, After 3 Years of Intervention). Using latent growth curve analysis, Wehmeyer and colleagues (2013) determined that students with cognitive disabilities who participated in interventions to promote self-determination over a three-year period showed significantly more positive patterns of growth in their self-determination scores than did students not exposed to interventions to promote self-determination.

Subsequently, in a follow-up study of the treatment and control group students from Wehmeyer et al. (2013), Shogren et al. (2015) investigated adult outcomes one and two years after leaving school. The study measured employment, community access, financial independence, independent living, and life satisfaction outcomes. Results indicated that self-determination status at the end of high school predicted significantly more positive employment, career goal, and community access outcomes. Students who were self-determined were significantly higher in all of these areas. These two studies provided causal evidence that promoting self-determination results in enhanced self-determination, and that enhanced self-determination results in more positive adult outcomes, including employment and community inclusion, and emphasize the need to get information about evidence-based practices into the hands of teachers.

One evidence-based intervention shown to improve student self-determination, access to the general education curriculum, and academic and transition goal attainment is the Self-Determined Learning Model of Instruction (SDLMI; Mithaug, Wehmeyer, Agran, Martin, & Palmer, 1998; Wehmeyer, Palmer, Agran, Mithaug, & Martin, 2000). The SDLMI is a model of instruction designed to enable teachers to support students to engage in self-regulated and self-directed learning. The SDLMI involves the use of self-regulated problem solving leading to the establishment of self-set goals, action plans to achieve those goals, and self-monitoring and self-evaluation activities to enable students to adjust plans and goals to attain the goal. The model was developed based on the component elements of self-determined behavior, the process of self-regulated problem solving, and research on student-directed learning (Wehmeyer et al., 2000). It is appropriate for use with students with and without disabilities across a wide range of content areas, and enables teachers to engage students in the totality of their educational program by increasing opportunities to self-direct learning and, in the process, to enhance student self-determination. Implementation of the model consists of a three-phase instructional process: Set a goal (Phase 1), Take action (Phase 2), and Adjust goal or plan (Phase 3). Each instructional phase presents a problem to be solved by the student (What is my goal? What is my plan? Have I achieved my goal?). The student solves each problem by posing and answering a series of four Student Questions per phase that students learn, modify to make their own, and apply to reach self-set goals. Each student question is linked to a set of Teacher Objectives that, in essence, describe the outcomes desired by having the student answer the question. Each instructional phase includes a list of Educational Supports that teachers can implement to enable students to self-direct learning. In each instructional phase, the student is the primary agent for choices, decisions, and actions, even when eventual actions are teacher-directed.

More than a dozen quasi-experimental or single-subject design studies have shown the potential efficacy of the SDLMI to promote self-determination and goal attainment (Wehmeyer, Abery, Mithaug, & Stancliffe, 2003). Two recent studies establish the SDLMI as an evidence-based practice. Wehmeyer et al. (2012) conducted a switching replication, randomized trial control group study on the impact of intervention with the SDLMI on student self-determination. Data on self-determination using multiple measures was collected
with 312 high school students with cognitive disabilities in both a control and treatment group. Wehmeyer and colleagues examined the relationship between the SDLMI and self-determination using structural equation modeling. After determining strong measurement invariance for each latent construct, these researchers found significant differences in latent means across measurement occasions and differential effects attributable to the SDLMI. This was true across disability category, though there was variance across disability populations. In other words, instruction using the SDLMI resulted in enhanced self-determination. Next, Shogren, Palmer, Wehmeyer, Williams-Diehm, and Little (2012) reported findings from a cluster or group randomized trial control group study examining the impact of the SDLMI on student academic and transition goal attainment and access to the general education curriculum for students with intellectual disability and learning disabilities. Students in the treatment group had significantly higher levels of goal attainment and access to the general education curriculum than their peers in the control group.

As noted previously, studies using SDLMI as an intervention have been conducted not only in the U.S. and Korea but also in other countries. In addition, there have been no meta-analytic studies of single-case design studies of the SDLMI in the U.S. or Korea. As a methodology, we opted to use the percentage non-overlapping data (PND) metric (Scruggs & Mastropieri, 1998; Scruggs Mastropieri, & Casto, 1987) to systematically investigate the overall effect of the SDLMI as an intervention with students with disabilities.

There is debate in the field as to the most appropriate metric to use in synthesizing single-case design research. Some researchers (Marquis et al., 2000) caution against the use of PND as an effect size measure for single subject research, suggesting it can be unduly affected by outliers in the baseline phase or by differing lengths in the various phases of the design. However, more recent studies that have compared findings for different single subject effect sizes, including PND, percentage zero data (PZD), standardized mean differences (SMD), mean baseline reduction (MBR), and regression based effect size estimates have reported that PND, when implemented correctly, is as successful as other effect size estimates in detecting intervention effects in single subject research meta-analyses (Campbell, 2000, 2004; Olive & Smith, 2005). In addition, PND has advantages over other effect size estimates. For example, regression-based effect size estimates of single-subject designs assume linearity, are complicated to calculate, and may, in some cases, overestimate the effect of treatment (Campbell, 2004). Further, Campbell, in a review of 117 studies that targeted reduction of problem behaviors of persons with autism found a significant correlation between PND and MBR, suggesting, as have the findings of other researchers (e.g., Olive & Smith, 2005), the comparability of PND, MBR, and SMD in detecting treatment effects. The results of a review of 68 syntheses using meta-analyses of single subject research for students with disabilities also support that PND was the most widely used as an effect size index, indicating that 47 effective sizes (54.65%) of a total of 84 effect sizes reported across 68 syntheses were calculated using the PND (Maggin, O’keeffe, & Johnson, 2011). In order words, in spite of continued criticism of the use of the PND in synthesizing single subject research, PND has contributed to provide

Method

Procedure

The study conducted a meta-analysis focused only on single-case research literature because the majority of SDLMI studies conducted in

Korea have used such designs, and because there have been no meta-analytic studies of single-case design studies of the SDLMI in the U.S. or Korea. As a methodology, we opted to use the percentage non-overlapping data (PND) metric (Scruggs & Mastropieri, 1998; Scruggs Mastropieri, & Casto, 1987) to systematically investigate the overall effect of the SDLMI as an intervention with students with disabilities.
coherent an valid syntheses of various research in a variety of subject areas and been considered as on one of the most versatile and meaningful methods (Scruggs & Mastropieri, 2013).

Based on recent research that suggests parametric approaches such as regression measures do not provide greater insight into effect sizes than do nonparametric approaches and, given the comparability of the PND metric with other estimates of effect size (MBR, SMD) (Campbell, 2004; Olive & Smith, 2005), the extensive history of using PND in the meta-analysis of single subject research designs, and the availability of standards for evaluating ceiling and outlier effects in the context of calculating PND (Scruggs et al., 1987; which were implemented in this study), the PND nonparametric approach was chosen in this study to synthesize the body of literature on SDLMI. Other variables that could potentially affect the efficacy of SDLMI interventions were also considered via examination of PND values across various participants and intervention characteristics.

Studies published prior to December 2012 that reported the effects of SDLMI interventions for students with disabilities were obtained through a search of several primary databases, including PsychINFO, ERIC, EBSCOhost, and Sage Publication along with a search of Korean database such as DBpia, RISS, and KISS. Various keywords and their combinations were used alone in combination for the search including Self-Determined Learning Model of Instruction (SDLMI), self-determination, and disability along with corresponding Korean keywords. Additional articles were found from the reference sections of the articles identified through the databases searches.

The following criteria were established to identify articles included in the analysis: (a) a single-case research design was employed; (b) the recipient of the intervention had a disability; (c) the SDLMI or a modified SDLMI version (i.e., the Self-Determined Career Development Model, SDCDM) were the primary intervention used to attain students’ target goals; (d) dependent variables were described clearly; (e) the effect of the intervention on the independent variables was empirically measured and graphically illustrated with clearly identifiable baseline and intervention phases; and (f) the article was published in a peer-reviewed journal.

Fifteen articles published between 2002 and 2012 were identified, with nine appearing across eight U.S. based journals (Remedial and Special Education, Career Development for Exceptional Individuals, Education and Training in Autism and Developmental Disabilities, Journal of Vocational Rehabilitation, Behavioral Disorders, Research and Practice for Persons with Severe Disabilities, Journal of Special Education, and Journal of Positive Behavior Interventions) and six being published across four Korean journals (Korean Journal of Special Education, Journal of Emotional and Behavioral Disorders, Journal of Special Education, and The Journal of Special Education: Theory and Practice). A total of 50 participants with disabilities were included in these studies. Thirty-two participants were male (mean age = 17.1) and 18 were female (mean age = 16.9).

Each study selected for analysis was coded relative to the following variables: (a) where the articles were published (i.e., U.S. and Korea); (b) the type of intervention; (c) the goals of intervention (target behaviors as dependent variables); (d) whether the study included maintenance and generalization data; and (e) the type of experimental research design. Analysis also indicated studies with two “levels” of the use of the SDLMI: in 73% (n = 11) of the studies the original SDLMI was used as the intervention; in 27% (n = 4) of the studies, a modified SDLMI was used, including the SDCDM or the SDLMI combined other instructional practices, such as Computer assisted instruction (CAI) or multimedia. In terms of the purpose of intervention related to dependent variables, two types of intervention goals were identified, the first related to access to general education curriculum (n = 9, 60%) and the second to transition-related outcomes such as career or employment preparation (n = 6, 40%). To investigate the direct effect of the SDLMI on dependent variables related to access to the general education curriculum or transition outcomes, other dependent variables, such as increasing knowledge of the SDLMI itself (Mazzotti, Test, et al., 2012; Mazzotti, Wood, Test, & Fowler, 2012) were excluded from the analysis process. All studies except one included maintenance or generalization infor-
information. Ten (67%) of the studies had maintenance information only and four (27%) studies included both maintenance and generalization data. Research designs included AB design \( (n = 1, 7\%) \), multiple baselines \( (n = 12, 80\%) \), and other \( (n = 2, 13\%) \).

Information about individual participants \( (n = 50) \) and context for the study were also coded: (a) age, (b) gender, (c) type of disability, and (d) setting in which the intervention took place. Ages of participants ranged from 10 to 50 (mean age 17.0). Students in elementary schools constituted 22\% \( (n = 11) \) of the sample; 50\% \( (n = 25) \) were middle and high school students and 28\% \( (n = 14) \) were youth and adult with disabilities. Males constituted 64\% of the sample and 36\% were female.

Several types of disability were identified as follows: intellectual disability (ID) \( (n = 29, 58\%) \), learning disability (LD) \( (n = 4, 8\%) \), Autism Spectrum Disorder (ASD) \( (n = 2, 4\%) \), Emotional and Behavior Disorder (EBD) \( (n = 5, 10\%) \), and other disabilities, such as multiple disabilities \( (n = 10, 20\%) \). The settings in which the interventions were implemented for each participant included: general education classrooms in public schools \( (n = 22, 44\%) \), separate classrooms \( (n = 5, 10\%) \); special schools \( (n = 11, 22\%) \); and community settings or workplaces \( (n = 12, 24\%) \).

**Analyses**

The percentage of non-overlapping data between treatment and baseline phases (Scruggs & Mastropieri, 1998; Scruggs et al., 1987) was calculated for the overall effect of intervention using the SDLMI on goals related either to increasing target behaviors or reducing target behaviors. For goals related to increasing target behaviors, PND was calculated by dividing the number of treatment data points that exceeded the highest baseline data point by the total number of data points in the treatment phase, and multiplying the sum by 100 (Scruggs et al., 1987). For goals related to reducing target behaviors, PND was calculated by dividing the number of treatment data points that fell below the lowest baseline data point by the total number of data points in the treatment phase, multiplied by 100 (Scruggs et al.). Thus, higher PND scores reflect more effective treatments (Scruggs et al.). To address concerns about the influence of outliers (Marquis et al., 2000), we followed the conventions described by Scruggs et al. (1987) to address ceiling or floor effects. Scruggs and Mastropieri (1998) suggested the following standards for evaluating PND values (scores can range from 0 to 100\%): scores below 50\% represent ineffectiive treatments; scores between 50\% and 70\% are questionable treatments; scores from 70\% to 90\% are effective treatments, and scores above 90\% represent highly effective treatments. For each article reviewed, PND scores were calculated for each unique treatment phase and its preceding baseline based on the graphs provided in each of the 15 studies. Descriptive statistics were also calculated for PND scores across unique intervention and participant variables.

Non-parametric procedures were used to analyze effects of various intervention and participant characteristics on PND scores (Scruggs et al, 1987; Scotti, Evan, Meyer, & Walker, 1991). The effects of country, intervention type, intervention goal, and gender were examined with separate Kolmogorov-Smirnov Two-Sample tests with average PND scores for each study and participant serving as the dependent measure. To investigate the effects of research design, and maintenance or generalization on treatment efficacy, separate Kruskal-Wallis One-Way ANOVAs were conducted using PNDs as dependent variables. In addition to intervention characteristics, the effects on treatment efficacy of participant’s age, disability type, and settings examined through separate Kruskal-Wallis One-Way ANOVAs (each participant’s mean PND score served as the dependent variable). Mann-Whiney test was conducted as follow-up test for the variables that showed statistical significance through Kruskal-Wallis One-Way ANOVA tests.

**Reliability**

The first author and a second rater independently calculated PND scores for each unique treatment phase and its preceding baseline to assess the reliability of PND scores. Agreement was scored when the author and second rater obtained identical PND scores for each phase. Reliability was calculated by dividing the number of agreements by the number of agree-
ments plus disagreements, multiplied by 100. Overall PND reliability was 95.3%. The first author then reviewed each of the articles, and coded the intervention and participant characteristics. The second rater independently coded intervention and participant characteristics for each of the articles as an additional reliability check. Interrater agreement for intervention and participant characteristics was calculated by dividing the number of agreements by the number of agreements plus disagreements, multiplied by 100. The overall interrater agreement for the intervention and participants characteristics was 94.6% and 95.5% respectively.

Results

PND analyses resulted in 64 unique PND scores. Given that some participant’s PND scores represented separate phases of treatment within the same participant due to research design and multiple goals, the average PND scores for each participant (and study) was calculated to control for the effects of individual characteristics (e.g., age, gender), without the confound of multiple phases per student. The overall mean PND score was 79.8% ($SD = 28.6$), with a range from 0 to 100%. Using Scruggs and Mastropieri’s (1998) criteria, this PND value (79.8%) represents an effective treatment. PND scores for intervention characteristics are presented in Table 1; mean PND scores for participant characteristics are presented in Table 2.

Table 3 presents the results of the Kolmogorov-Smirnov Two-Sample tests of intervention and participant characteristics. The only significant effect was that of intervention type (Kolmogorov-Smirnov $Z$ score = 1.56, $p = .02$). The mean PND score for original SDLMI was 89.7% while the mean score for modified SDLMI was 67.6%, as described in Table 1, meaning that the former was more effective in meeting the intervention goals than did the latter. The results of the Kruskal-Wallis One Way ANOVA of intervention and participant characteristics revealed that there was only significant effect in the age of participants ($\chi^2(2, N = 50) = 16.58, p = .00$), as presented in Table 4.

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean PND (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S</td>
<td>9</td>
<td>81.9 (19.1)</td>
</tr>
<tr>
<td>Korea</td>
<td>6</td>
<td>86.6 (28.2)</td>
</tr>
<tr>
<td>Intervention Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDLMI</td>
<td>11</td>
<td>89.7 (20.9)</td>
</tr>
<tr>
<td>Modified SDLMI</td>
<td>4</td>
<td>67.6 (19.7)</td>
</tr>
<tr>
<td>Intervention Goal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access to the General Education Curriculum</td>
<td>9</td>
<td>81.7 (27.8)</td>
</tr>
<tr>
<td>Transition-related Outcomes</td>
<td>6</td>
<td>87.0 (11.7)</td>
</tr>
<tr>
<td>Maintenance/Generalization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>70.0 (N/A)</td>
</tr>
<tr>
<td>Maintenance</td>
<td>10</td>
<td>91.9 (8.3)</td>
</tr>
<tr>
<td>Maintenance+Generalization</td>
<td>4</td>
<td>67.0 (38.3)</td>
</tr>
<tr>
<td>Research Design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td>1</td>
<td>78.6 (N/A)</td>
</tr>
<tr>
<td>Multiple Baseline</td>
<td>12</td>
<td>87.4 (17.9)</td>
</tr>
<tr>
<td>Mixed</td>
<td>2</td>
<td>64.6 (50.1)</td>
</tr>
</tbody>
</table>

Discussion

These findings provide additional support for the efficacy and benefit of SDLMI as an intervention teachers can implement to support students with disabilities to reach goals related to academic and functional areas. That is, across countries, participants, and settings, intervention with the SDLMI resulted in improvement in academic and functional outcomes related to access to the general education curriculum and preparing students for a successful transition to adulthood (mean PND score = 79.8%). The average PND value of 79.8% indicates that use of SDLMI as an intervention is generally an effective way to promote access to the general education curriculum and transition outcomes along with enhancing problem-solving and goal setting and attainment. As with other intervention strategies for students with disabilities, however, the SDLMI will neither be universally effective nor suitable for all students and their needs. Nevertheless, these results provide additional support to group randomized trial studies that the SDLMI is effective as an intervention.
The Impact of Intervention and Participant Characteristics

In terms of the analysis of intervention and participants characteristics, no variables reached statistical significance on the Kolmogorov-Smirnov Two-Sample and Kruskal-Wallis One Way ANOVA tests, except intervention type and participants’ age. However, the trends in the PND scores across the intervention and participant characteristics provide preliminary comparative and speculative opportunities and prospects for future investigation.

**Intervention characteristics.** The effects of SDLMI were clearly non-significant between countries, although the mean PND scores ($M = 86.6\%$) of Korean studies were a little higher than that of studies conducted in the U.S. ($M = 81.9\%$). Given that the mean PND scores of studies in both the U.S. and Korea were over 80%, SDLMI can be considered an effective intervention in both countries. This suggests that the SDLMI could be applied to promote student goal attainment in other countries. To this end, future research is needed to investigate the effects of the SDLMI in other countries by applying SDLMI in various countries and reviewing more studies of SDLMI conducted in other countries.

Interventions using the SDLMI were divided into two types; original SDLMI and modified versions of the SDLMI, including the Self-Determined Career Developmental Model (SDDCM)—which was based on identical framework and process of the SDLMI, but focuses solely on career and employment goals for adults with disabilities (Benitez et al., 2005; Wehmeyer et al., 2003)—and the SDLMI combined with other instructional strategies, such as Computer-assisted instruction (CAI) and Multimedia Goal-Setting Instruction (MGSI) (Mazzotti, Test, et al., 2012; Mazzotti, Wood, et al., 2012). The studies using the original version of the SDLMI ($n = 11$) were

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean PND (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary School Students</td>
<td>11</td>
<td>55.9 (26.2)</td>
</tr>
<tr>
<td>Middle &amp; High School Students</td>
<td>25</td>
<td>91.1 (13.4)</td>
</tr>
<tr>
<td>Youth &amp; Adults</td>
<td>14</td>
<td>87.2 (20.8)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>32</td>
<td>80.4 (26.4)</td>
</tr>
<tr>
<td>Female</td>
<td>18</td>
<td>85.6 (16.9)</td>
</tr>
<tr>
<td>Disability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intellectual Disability</td>
<td>29</td>
<td>87.6 (18.0)</td>
</tr>
<tr>
<td>Learning Disability</td>
<td>4</td>
<td>86.6 (15.5)</td>
</tr>
<tr>
<td>Autism Spectrum Disorder</td>
<td>2</td>
<td>100 (.00)</td>
</tr>
<tr>
<td>Emotional &amp; Behavior Disorder</td>
<td>5</td>
<td>84.3 (15.1)</td>
</tr>
<tr>
<td>Others</td>
<td>10</td>
<td>60.6 (33.0)</td>
</tr>
<tr>
<td>Setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Education Classroom</td>
<td>22</td>
<td>86.9 (19.1)</td>
</tr>
<tr>
<td>Separate Classroom</td>
<td>5</td>
<td>51.2 (38.0)</td>
</tr>
<tr>
<td>Special School</td>
<td>11</td>
<td>88.9 (13.3)</td>
</tr>
<tr>
<td>Community/Workplace</td>
<td>12</td>
<td>80.8 (22.9)</td>
</tr>
</tbody>
</table>

**TABLE 3**

<table>
<thead>
<tr>
<th>Kolmogorove-Smirnov Results for PND scores of Intervention and Participant Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Intervention Characteristics</td>
</tr>
<tr>
<td>Country</td>
</tr>
<tr>
<td>Intervention Type</td>
</tr>
<tr>
<td>Intervention Goal</td>
</tr>
<tr>
<td>Participant Characteristics</td>
</tr>
<tr>
<td>Gender</td>
</tr>
</tbody>
</table>

---

**TABLE 4**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$X^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention Characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance/Generalization</td>
<td>1.75</td>
<td>.42</td>
</tr>
<tr>
<td>Research Design</td>
<td>1.09</td>
<td>.58</td>
</tr>
<tr>
<td>Participant Characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>16.58</td>
<td>.00</td>
</tr>
<tr>
<td>Setting</td>
<td>6.08</td>
<td>.11</td>
</tr>
<tr>
<td>Disability</td>
<td>8.57</td>
<td>.08</td>
</tr>
</tbody>
</table>

*df = 2 for each analysis
*bdf = 3 for each analysis
*cdf = 4 for each analysis

SDLMI for Students with Disabilities / 243
almost three times more numerous than studies of modified versions \((n = 4)\), and interventions using the original version of the SDLMI resulted in higher PND scores \((M = 89.7\%)\) than did interventions with modified versions \((M = 67.6\%)\). These differences were statistically significant (Kolmogorov-Smirnov \(Z = 1.56, p = .02\)). In other words, intervening using the original version of the SDLMI was more effective that using modified versions. There have, however, been few examinations of the SDCDM, and the jury remains out with regard to its full efficacy.

As dependent variables, intervention goals targeted through the SDLMI were related to access to the general education curriculum and transition-related outcomes. Goals pertaining to access to the general education curriculum included promoting problem-solving skills, active classroom participation skills, self-regulated learning strategies, academic achievements, self-expressions, requesting for help from others, and reducing disruptive behaviors in general education classrooms. In terms of transition goals related to career and employment, several target behaviors were included such as promoting student involvement in career and vocational planning and decision-making, enhancing job performances, and working on the goals for obtaining specific job skills. The mean PND scores \((M = 87.0\%)\) of intervention goals for transition-related outcomes was higher than that for access to the general education curriculum \((M = 81.7\%)\), but this difference was not statistically significant. The mean PND scores of both intervention goal areas being over 80% reflects the potential of the SDLMI as an effective intervention to enhance both access to the general education curriculum and transition outcomes of students with disabilities.

**Participant characteristics.** The mean PND values varied depending on participants’ age groups. Middle and high school students constituted 50\% \((n = 25)\) of all participants and the mean PND score for this group \((M = 91.1\%)\) was highest compared with other participants, followed by that of older youth and adults \((M = 87.2\%)\). There were only 11 elementary school students who participated in studies in which the SDLMI was implemented, and the mean PND score for that group was lowest \((M = 55.9\%)\). The results of the Kruskal-Wallis One Way ANOVA showed that the differences were statistically significant \((\chi^2(2, N = 50) = 16.58, p = .00)\). Follow-up tests using Mann-Whitney test were conducted to evaluate pairwise differences among the three age groups, controlling for Type I error across tests by using the Bonferroni approach. The results of these tests indicated a significant difference not only between elementary school students and middle and high school students \((Z = -3.960, p < .017)\), but also between elementary school students and youth and adults \((Z = -2.962, p < .017)\). In other words, the effect of SDLMI as an intervention was statistically more effective for adolescents, youth and adults with disabilities. Again, though, it seems unwise to generalize this finding at this point, given that the studies analyzed included significantly more middle and high students and relatively few elementary students. But, it is also logical to assume that younger students may benefit from such instruction, though perhaps not at the same level as older students, who have more of the skills necessary to self-regulate learning. Certainly, more research is needed to investigate the impact of instruction using the SDLMI with elementary school students. Female’s average PND scores \((M = 85.6\%)\) tended to be higher than male’s scores \((M = 80.4\%)\), although, again, this difference did not reach statistical significance.

In terms of types of disability, intellectual disability was the most frequently identified disability category \((n = 29, 58\%)\) and Autism Spectrum Disorder (ASD) was the least found \((n = 2, 4\%)\). However, the mean PND score \((M = 100\%)\) for ASD was the highest, followed by intellectual disability \((M = 87.6\%)\), learning disability \((M = 86.6\%)\), emotional and behavior disorder \((M = 84.3\%)\), and other disabilities \((60.6\%)\). Although the differences were not statistically significant, future study is need in this area given that the number of sample across disability categories was clearly different and majority of disability were intellectual disability.

This study provided evidence of the potential utility of the SDLMI across various settings, given that the mean PND scores were relatively high across different settings such as general education classrooms \((M = 86.9\%)\), special schools \((M = 88.9\%)\) and communi-
ties or workplaces \((M = 80.8\%)\), except separate classrooms \((M = 51.2\%)\). Although these differences were not statistically significant, the lowest mean PND score for separate classroom may indicate that SDLMI was effective when it was used in more natural settings where participants’ mainly study or work. Future investigation is need in this area as well given that the small number of studies conducted in separate classrooms might affect the differences.

**Limitations**

There were several limitations that should be considered in evaluating implications of the present study. First, as mentioned before, the study used a small sample size due to a limited number of studies that met our criteria. Second, there was no attempt to evaluate and judge the quality of the articles included in the analyses. Therefore, any methodological issues or errors in the included studies could have influenced the analysis process of the present study. Last, the studies reviewed were published only in the U.S and Korea, so there was limitation not to consider potential effects of SDLMI that might be applied in other countries.

**Conclusion**

Promoting self-determination is best practice in special education and research shows a causal relationship between promoting self-determination and attainment of more positive adult and school-related outcomes, including the attainment of academic and transition-related goals. This study provides further evidence, by performing a meta-analysis of single-case design studies of the SDLMI across two countries, of the interventions’ efficacy as a means to promote academic and functional goal attainment.

**References**

*References marked with an asterisk indicate studies included in the meta-analysis.

*Agran, M., Blanchard, C., Wehmeyer, M., & Hughes, C. (2002). Increasing the problem-solving skills of students with developmental disabili-


Received: 30 October 2013
Initial Acceptance: 6 January 2014
Final Acceptance: 10 June 2014
Friendship 101: Helping Students Build Social Competence

The eighth volume of the CEC Division on Autism and Developmental Disabilities’ Prism series, *Friendship 101* focuses on building social competence, friendship making, and recreation and leisure skills among students with autism spectrum disorder and other developmental disabilities. Chapters in this evidence-based, user-friendly guide address the needs of students in different developmental periods (from pre-K through young adulthood), providing teachers, parents, faculty and teacher educators with tools and strategies for enhancing the social skill development of these children and youth. Presented through an ecological perspective, together these chapters emphasize building social competence within and across school, home, and community contexts.

Available from CEC Publications
https://www.cec.sped.org/Publications
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.
17th International Conference on Autism, Intellectual Disability & Developmental Disabilities

Research-Informed Practice

January 20 - 22, 2016

Waikiki Beach, Hawaii

Council for Exceptional Children Division on Autism and Developmental Disabilities

CEC’s Division on Autism and Developmental Disabilities (DADD) is hosting the 17th International Conference, which is both research and practitioner-focused. DADD is pleased to extend an invitation to join us in Waikiki Beach, Hawaii, January 20 - 22, 2016, for a stellar professional learning opportunity! The program will feature more than 100 lecture and poster presentations in the following topical areas:

- Autism Spectrum Disorder
- Assistive & Adaptive Technology
- Early Childhood
- Intellectual Disability
- Post-Secondary Initiatives
- Multiple Disabilities
- Paraprofessionals
- Dual Diagnosis
- Mental Health
- Transitions

For further information, please contact:

Cindy Perras, Conference Co-ordinator
cindy.perras@cogeco.ca
www.daddcec.org