Education and Training in Developmental Disabilities

Focusing on individuals with cognitive disabilities/mental retardation, autism, and related disabilities

Volume 41
Number 3
September 2006
Education and Training in Developmental Disabilities
The Journal of the Division on Developmental Disabilities,
The Council for Exceptional Children

Editor: Stanley H. Zucker
Arizona State University

Editorial Assistant: Amy Barry
Arizona State University

Consulting Editors
Martin Agran
Reuben Altman
Phillip J. Bellmore
Sharon Borthwick-Duffy
Michael P. Brady
Fredda Brown
Mary Lynne Calhoun
Sharon F. Cramer
Caroline Dunn
Lise Fox

Consulting Editors
David L. Gast
Herbert Goldstein
Carolyn Hughes
Larry K. Irvin
James V. Kahn
H. Earle Knowlton
Barry W. Lavay
Rena Lewis
Kathleen J. Marshall

Advertising rates are available upon request.

Manuscripts should be typed, double spaced, and sent (five copies) to the Editor: Stanley H. Zucker, Special Education Program, Box 872011, Arizona State University, Tempe, AZ 85287-2011. 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, 2001 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.


Division on Developmental Disabilities
The Council for Exceptional Children

Board of Directors

Officers
Past President Dagny Fidler
President Phil Parette
President-Elect Polly Parrish
Vice President J. David Smith
Secretary Cindy Perris
Treasurer Amanda Boutot

Members
Leslie Broun
Jim Forristal
Kent Gerlach
Kara Hume (Student Governor)
Nikki Murdick
Dianne Zager

Executive Director
Tom E. C. Smith
Publications Chair
Jack Hourcade
Communications Chair
Darlene Perner

The purposes of this organization shall be to advance the education and welfare of persons with developmental disabilities, research in the education of persons with developmental disabilities, competency of educators in this field, public understanding of 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 DEVELOPMENTAL DISABILITIES (ISSN 1547-0350) (USPS 0168-5000) is published quarterly in March, June, September, and December, by The Council for Exceptional Children, Division on Developmental Disabilities, 1110 North Glebe Road, Arlington, Virginia 22201-5704. Members’ dues to The Council for Exceptional Children Division on Developmental Disabilities include $8.00 for subscription to EDUCATION AND TRAINING IN DEVELOPMENTAL DISABILITIES. Subscription to EDUCATION AND TRAINING IN DEVELOPMENTAL DISABILITIES is available without membership; Individual—U.S. $40.00 per year; Canada, PUAS, and all other countries $44.00; Institutions—U.S. $100.00 per year; Canada, PUAS, and all other countries $104.50; single copy price is $25.00. U.S. Periodicals postage is paid at Arlington, Virginia 22204 and additional mailing offices.

POSTMASTERS: Send address changes to EDUCATION AND TRAINING IN DEVELOPMENTAL DISABILITIES, 1110 North Glebe Road, Arlington, Virginia 22201-5704.
Education and Training in Developmental Disabilities

Editorial Policy

*Education and Training in Developmental Disabilities* focuses on the education and welfare of persons with developmental disabilities. *ETDD* invites research and expository manuscripts and critical review of the literature. Major emphasis is on identification and assessment, educational programming, characteristics, training of instructional personnel, habilitation, prevention, community understanding and provisions, and legislation.

Each manuscript is evaluated anonymously by three reviewers. Criteria for acceptance include the following: relevance, reader interest, quality, applicability, contribution to the field, and economy and smoothness of expression. The review process requires two to four months.

Viewpoints expressed are those of the authors and do not necessarily conform to positions of the editors or of the officers of the Division.

Submission of Manuscripts

1. Manuscript submission is a representation that the manuscript is the author’s own work, has not been published, and is not currently under consideration for publication elsewhere.
2. Manuscripts must be prepared according to the recommendations in the *Publication Manual of the American Psychological Association* (Fifth Edition, 2001). Standard typewriter type, laser, or high density dot printing are acceptable.
3. Each manuscript must have a cover sheet giving the names and affiliations of all authors and the address of the principal author.
4. Graphs and figures should be originals or sharp, high quality photographic prints suitable, if necessary, for a 50% reduction in size.
5. Five copies of the manuscript along with a transmittal letter should be sent to the Editor: Stanley H. Zucker, Special Education Program, Box 872011, Arizona State University, Tempe, AZ 85287-2011.
6. 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.
7. The Editor reserves the right to make minor editorial changes which do not materially affect the meaning of the text.
8. Manuscripts are the property of ETDD for a minimum period of six months. All articles accepted for publication are copyrighted in the name of the Division on Developmental Disabilities.
Curriculum Augmentation and Adaptation Strategies to Promote Access to the General Curriculum for Students with Intellectual and Developmental Disabilities

SUK-HYANG LEE
BETTY A. AMOS
STELIOS GRAGOUDAS
YOUNGSUN LEE
KARRIE A. SHOGREN
RASCHELLE THEOHARIS
MICHAEL L. WEHMEYER

Functional Communication Training: A Review of the Literature Related to Children with Autism

G. RICHMOND MANCIL

Effectiveness of Simultaneous Prompting in Small Group: The Opportunity of Acquiring Non-target Skills through Observational Learning and Instructive Feedback

OGUZ GURSEL
ELIF TEKIN-IFTAR
FUNDA BOZKURT

Increasing Opportunities for Requesting in Children with Developmental Disabilities Residing in Group Homes through Pyramidal Training

RALF W. SCHLOSSER
ELIZABETH WALKER
JEFF SIGAFOOS

Computer-Based Instruction for Purchasing Skills

KEVIN M. AYRES
JOHN LANGONE
RICHARD T. BOON
AUDREY NORMAN

Increasing the Skills of Children with Developmental Disabilities through Staff Training in Behavioral Teaching Techniques

ANNA-LIND PETURSDÓTTIR
ZUILMA GABRIELA SIGURDARDÓTTIR

High School Teachers’ Perceptions of School-to-Work Transition Practices in Taiwan

MENG-CHI CHAN
JANIS G. CHADSEY
Reliability and Validity of the *Pervasive Developmental Disorders Rating Scale* and the *Gilliam Autism Rating Scale*

RONALD C. EAVES
SUZANNE WOODS-GROVES
THOMAS O. WILLIAMS JR.
ANNA-MARIA FALL

---

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

**December 2006**

Effect of adapted “Cover Write” method to teach spelling to students with developmental disabilities. **Dilek Erbas**, Arzu Ozen, Yasemin Turan, and James W. Halle, Erciyes University, Egitim Fakultesi, Kayseri, TURKEY.

Relationship of muscular strength on work performance in high school students with mental retardation. **Karen Smail** and Michael Horvat, College of Charleston, Physical Education and Health Department, 66 George St., Charleston, SC 29424-0001.

Using systematic instruction to teach decoding skills to middle school students with moderate intellectual disabilities. Stacey Bradford, Margaret E. Shippen, Paul Alberto, **David E. Houchins**, and Margaret Flores, Georgia State University, Department of Educational Psychology and Special Education, College of Education, P.O. Box 3979, Atlanta, GA 30302-3979.

Building math fluency for students with developmental disabilities and attention difficulties using Great Leaps Math. **Kristine Jolivette**, Amy S. Lingo, David E. Houchins, Sally Barton-Arwood, and Margaret E. Shippen, Educational Psychology and Special Education, P.O. Box 3979, Georgia State University, Atlanta, GA 30302-3979.

Comparison of three video-based instructional procedures for teaching daily living skills to persons with developmental disabilities. **Toni Van Laarhoven** and Traci Van Laarhoven-Myers, Department of Teaching & Learning, Northern Illinois University, DeKalb, IL 60115-2854.

Pathfinding in the research forest: The pearl harvesting method for effective information retrieval. **Robert Sandieson**, Faculty of Education, University of Western Ontario, 1137 Western Road, London, Ontario, N6G 1G7 CANADA.

Comparing video prompting to video modeling for teaching daily living skills to six adults with developmental disabilities. **Helen Cannella**, Jeff Sigafoos, Mark O’Reilly, Berenice de la Cruz, Chaturi Edrisinha, and Giulio E. Lancioni, The Ohio State University, School of Physical Activity & Educational Services, 356 Arps Hall, 1945 North High Street Columbus, OH 43210.


Use of response cards to teach telling time to students with moderate and severe disabilities. Channon Horn, **John W. Schuster**, and Belva C. Collins, 229 Taylor Building, Department of Special Education and Rehabilitation Counseling, University of Kentucky, Lexington, KY 40506-001.

Address is supplied for author in boldface type.
Curriculum Augmentation and Adaptation Strategies to Promote Access to the General Curriculum for Students with Intellectual and Developmental Disabilities

Suk-Hyang Lee, Betty A. Amos, Stelios Gragoudas, Youngsun Lee, Karrie A. Shogren, Raschelle Theoharis, and Michael L. Wehmeyer
University of Kansas

Abstract: Curriculum modification strategies, particularly curriculum adaptations and augmentations, have been identified as important to enable learners with disabilities to achieve access to and progress in the general curriculum. There is, however, relatively little research on the effect of these strategies with students with intellectual and developmental disabilities. The purpose of this paper is to examine curriculum adaptation and augmentation strategies that might promote student involvement and progress in the general curriculum for students with intellectual and developmental disabilities, and to provide recommendations with regard to how such augmentations might be modified to be more appropriate for use with this population.

More than five years after the 1997 amendments to IDEA first required that the IEPs of all students receiving special education services describe how a child’s disability affects his or her involvement with and progress in the general curriculum and provide statements concerning measurable goals, services, and program modifications to achieve such involvement and progress, there are still too few frameworks that describe strategies to address “access to the general curriculum” for all students with disabilities, particularly students with intellectual and developmental disabilities. Many educators working with students with more severe disabilities are dubious that the focus on access is either achievable or advisable. Agran, Alper, and Wehmeyer (2002) conducted a survey of teachers working with students with severe disabilities about their perceptions of the IDEA access requirements. When asked if ensuring students’ access to the general curriculum would help increase educational expectations for students with severe disabilities, 75% of teachers agreed to some degree. However, 63% indicated they felt access to the general education curriculum was more important for students with mild disabilities. While between 11% and 23% of respondents indicated they used several different ways to ensure some level of access, the largest proportion (37%) indicated that students with severe disabilities were receiving an educational program developed outside the context of the general curriculum. Nearly ¾ of respondents indicated that their students with severe disabilities were evaluated exclusively by criteria stipulated in the IEP. The majority of teachers (85%) indicated that students with severe disabilities should not be held to the same standards as students without disabilities, and over half (53%) reported their school district had no clear plan for ensuring access to the general curriculum for students with severe disabilities.

Such skepticism might, rightfully, be linked to the lack of concrete strategies forwarded to enable students with more severe disabilities to access the general curriculum. The lack of focus on access for this population is, however, slowly diminishing. Researchers and policymakers have proposed models to promote access for this population (Janney & Snell, 2000; Wehmeyer, Lance, & Bashinski, 2002), addressed issues concerning how to ensure an appropriate education within the
context of the access mandates (Browder, Flowers, Ahlgrim-Delzell, Karvonen, Spooner, & Algozzine, 2004; Kochhar-Bryant, & Bassett, 2002; Wehmeyer, Field, Doren, Jones, & Mason, 2004; Wehmeyer, Sands, Knowlton, & Kozleski, 2002), and conducted research examining factors contributing to access for students with intellectual and developmental disabilities (Palmer, Wehmeyer, Gipson, & Agran, 2004; Wehmeyer, Lattin, Lapp-Rincker, & Agran, 2003).

A topic of virtually all discussions concerning access for students with disabilities involves the centrality of curriculum modifications to this effort (Fisher & Nancy, 2001; Janney & Snell, 2000; Kame’enui & Simmons, 1999; Nolet & McLaughlin; 2000; Wehmeyer, Sands et al., 2002). Wehmeyer and colleagues (Knowlton, 1998; Wehmeyer, Lance et al., 2002; Wehmeyer, Lattin, & Agran, 2001; Wehmeyer, Sands et al., 2002) proposed a multi-level model to promote access for students with intellectual disabilities that involved three levels of curriculum modification: curriculum adaptations, augmentations, and alterations.

Curriculum adaptations refer to efforts to modify the way in which content is represented or presented or in which the student engages with and responds to the curriculum, including the implementation of features of Universal Design for Learning (Rose & Meyer, 2002; Wehmeyer, Lance et al., 2002). Curriculum augmentations refer to efforts to augment or expand the curriculum to provide students with additional skills or strategies that enable them succeed within the general curriculum. Curriculum alterations refer to the addition of content specific to a student’s needs, including functional skills or life skills not found in the general curriculum.

In observational studies of the degree to which these levels of curriculum modification were in place to support students with intellectual disabilities, Wehmeyer, Lattin et al. (2003) and Soukup, Wehmeyer, Bashinski, and Bovaird (2004) found that a few instances of curriculum adaptations were implemented, while no instances of the use of curriculum augmentations with this population were noted. These curriculum modification strategies are fundamental in efforts to promote progress in the general curriculum for students with other disabilities, particularly, students with learning disabilities (Deshler, Schumaker, Harris, & Graham, 1999; Kame’enui & Simmons, 1999; Lenz, Deshler, & Kissam, 2003; Scruggs & Mastropieri, 2000). Most of the focus in curriculum adaptations for students with more severe disabilities has been on the promising role technology can play in providing universally-designed materials (Rose & Meyer, 2002; Wehmeyer, Lance et al., 2002). There has, however, been relatively little research on potential curriculum adaptation and augmentation strategies that might support students with intellectual and developmental disabilities.

The purpose of this paper is to examine curriculum adaptation and augmentation strategies that might promote involvement and progress in the general curriculum for students with intellectual and developmental disabilities, and to provide recommendations with regard to how such adaptations and augmentations might be modified to be more appropriate for use with this population.

Curriculum Augmentation and Adaptation Strategies

Learning Strategies

Cognitive or learning strategies provide students with strategies that enable them to engage the learning process more effectively (Rosenthal-Malek & Bloom, 1998). There are a variety of such learning strategies (including shadowing, verbatim notes, graphic or advance organizers, semantic maps, mnemonics, chunking, questioning, and visualizing strategies), that fall under the category of curriculum adaptations or augmentations, all of which have been validated with students with learning disabilities and some of which might benefit students with intellectual or developmental disabilities (Rosenthal-Malek & Bloom). The following brief overview highlights those strategies that might warrant closer scrutiny with this population.

Graphic Organizers

Graphic organizers are “visual displays teachers use to organize information in a manner
that makes the information easier to understand and learn” (Meyen, Vergason, & Whelan, 1996, p. 132). They involve efforts to structure information or arrange important aspects of a concept or topic into a pattern using graphic modalities (Bromley, Irwin-DeVitis, & Modlo, 1995), and thus are curriculum adaptations (e.g., modifying content representation or presentation). Graphic organizers are effective in enabling students to assimilate new information by organizing previous information. Flow charts, semantic maps, webs, and Venn diagrams are all examples of graphic organizers.

A number of studies with students with learning disabilities have validated the efficacy of graphic organizers in improving text and reading comprehension (Alvermann, Boothby, & Woolf, 1984; Barron & Schwartz, 1984; Bos & Anders, 1992; Griffin, Simmons, & Kame’enui, 1991; Moore & Readence, 1980; Simmons, Griffin, & Kame’enui, 1988). In addition, graphic organizers have been effectively applied across other content areas, such as science, math, and social studies (Anstrong, 1993; Griffin et al.; Guastello, Beasley, & Sinatra, 2000; Hanselman, 1996). Visually displaying key content ideas can especially benefit students who struggle with organizing information (Fisher & Shumaker, 1995). In addition to improving student learning, graphic organizers have been shown to be useful in building relationships between students by sharing personal information such as hobbies, dreams, family and experiences with other students through “me maps” (Cullinan, Galda, & Strickland, 1993). Graphic organizers also can be applied at the whole-class level (Baxendell, 2003).

The limited number of research studies on the efficacy of the use of advance (though not specifically graphic) organizers with students with intellectual disabilities provided mixed results. Peleg and Moore (1982) found that when students with mild mental retardation were instructed using an advance organizer (either oral or written), the oral organizer seemed detrimental to learning while the written organizer led to a higher mean questions answered, although the latter did not reach significance. Subsequent research was more encouraging. Reis (1986) found that advance organizers in the form of knowledge statements (defines certain concepts in the content in advance), and purposive statements (provides students with a description of what he or she was supposed to listen for in particular) improved comprehension performance of students with and without intellectual disabilities (group), with all students performing better in the knowledge plus purpose statements condition than in all other conditions (knowledge statement only, purpose statement only, no advance organizer). Both the knowledge statement only and the purpose statement only conditions, however, were more positive than the no advance organizer condition. There were group differences in comprehension scores (e.g., students without intellectual disabilities answered, on the average, more questions than students with intellectual disabilities), but there were no group by treatment effects, indicating that students with intellectual disabilities received equivalent benefit from using the advance organizer. Similarly, Chang (1986) found that the use of an advance organizer prior to viewing a film facilitated comprehension for students with and without intellectual disabilities, with no differential effect based on disability (e.g., students with intellectual disabilities benefited as much from the advance organizers as students without intellectual disabilities).

The limited number of studies examining the potential utility of graphic or advance organizers for students with intellectual disabilities provides only limited information about their viability with this population, but given this strategy’s prominence in the field of learning disabilities as an effective way to adapt the representation and presentation of curriculum content, it is important to consider this approach more seriously. Moreover, there have not been extensive efforts to examine what types of graphic organizers might be effective for this population (other than to suggest that oral or verbal organizers may not be the best means) and how modifications to more traditional graphic or advance organizers might have efficacy for students with intellectual and developmental disabilities. Specifically, the use of computer-based technologies provides newer and potentially more powerful ways, through features such as multimedia presentation of ‘big ideas,’ to provide graphic organizers for all students and particularly for students with intellectual disabilities.
**Chunking**

The definition and strategies for chunking vary somewhat depending on the content areas or contexts in which this strategy is used. However, chunking is basically the process of “combining related elements into units” (Sylwester, 1995) that are manageable to students. Chunking is a curriculum augmentation strategy in that students learn to ‘chunk’ material to make it more manageable and to improve memory and recall. Chunking has been used as a teaching device in content enhancement for students with learning disabilities (Bulgren & Lenz, 1996). Chunking is especially effective in improving skills related to language arts, such as reading (Cortese, 2003; Silliman, Bahr, Beasman, & Wilkinson, 2000; Vogt & Nagan, 2003), word recognition (Morris, Bloodgood, Lomax, & Perney, 2003), verbal recall (Montgomery, 2002), spelling (by chunking letter and matching sounds) (Dahl et al., 2003), fluency and comprehension (Vaughn et al., 2000) and writing (Short, Kane, & Peeling, 2000). In addition to its efficacy with students with learning disabilities, chunking has been shown to be effective in improving word analysis, reading, and recall information for students who are English language learners, students with attention-deficit-hyperactivity disorder, and gifted students (Gallagher, 1994; Linan-Thompson, Vaughn, Hickman-Davis, & Kouzekanani, 2003; Schwiebert, Sealandier, & Dennison, 2002). Moore and Brantingham (2003) taught a student with reading difficulties to study his own miscues to improve his reading through Retrospective Mis-cue Analysis (RAM), which incorporated chunking as one strategy. Sentence by sentence self-monitoring (SSSM), which included a chunking strategy, was also effective in enabling students to be active readers by internalizing self-monitoring (Buettnert, 2002). In addition, chunking was used to help students with academic and behavior problems succeed in school by enabling them to learn social and self-management skills, including specific cognitive skills such as “chunking key ideas into small groups” (Brigman & Campbell, 2003). Across virtually all of these studies, direct instruction to teach the student the chunking strategy is necessary for the students to utilize and benefit from the strategy (Short et al., 2000).

While there is much research that focuses on the effect of chunking for students with learning disabilities, struggling readers, and young emergent readers, there is little research for students with intellectual disabilities. One outcome of standards-based reform efforts through the No Child Left Behind Act and the IDEA Access to the General Curriculum mandates has been to focus more attention on teaching students with intellectual disabilities to read (Browder & Spooner, in press). Based on this strategy’s utility in improving reading and language arts outcomes for students with other disabilities and given the strategy’s close links with self-management strategies, which have been shown to be effective with this population, it would seem worthwhile to examine the efficacy of this strategy with students with intellectual and developmental disabilities. Given the well-documented difficulty students with intellectual and developmental disabilities have with memory, chunking might be a useful means to enhance attention and memory by enabling them to chunk related parts into units that are meaningful to them. Combining chunking strategies with other strategies, such as graphic organizers, and using the capacity of technology may make the strategy more accessible for students with intellectual and developmental disabilities. Using chunking with other visual stimuli or cues could facilitate a student’s attention to and memory about specific chunks. In addition, chunking could be a bridge between teacher-directed teaching and student-directed learning strategies by gradually fading teachers’ direct instruction and empowering students to be more active learners.

**Mnemonic Strategies**

Mnemonic strategies are systematic procedures for enhancing memory by providing effective cues for recall as a “cognitive cuing structure” such as word, sentence, or picture devices (Bellezza, 1981; Lombardi & Butera, 1998). This strategy is used mainly in developing better ways to encode new information for easier retrieval (Mastropieri & Scruggs, 1998). Mnemonic strategies are commonly divided into imagery illustrations, such as pictures or
diagrams, and word-based devices, using words to aid memory (Access Center, 2003; Mastropieri & Scruggs, 1991; Daniel & Pressley, 1987). It is more effective, however, to integrate imagery illustration and word-based devices as opposed to using them separately. Many studies have documented the efficacy of mnemonic strategies with students with learning disabilities and mild intellectual disabilities. Such strategies can: be used across multiple content areas, such as language arts, mathematics, science and social studies (Mastropieri & Scruggs, 1988, 1991; Scruggs & Mastropieri, 2000); be used across age ranges, from first-grade to adolescence or adulthood (Fulk, Lohman, & Belfiore, 1997; Scruggs & Mastropieri); be used for behavioral interventions, including self-management and positive behavior support (Agran, King-Sears, Wehmeyer, & Copeland, 2003; Silverstein, 1997; Smith, Siegel, O’Connor, & Thomas, 1994); be a tool for other curriculum modification strategies such as cognitive mapping, computer-assisted instruction, and self-monitoring or self-instruction (Boyle & Yeager, 1997; Brown & Frank, 1990; Irish, 2002); and can be effectively applied to the classroom level for access to the general curriculum (Ashton, 1999; Mastropieri, Scruggs, & Whedon, 1997; Mastropieri & Scruggs, 1998; Munk, Bruckert, Call, Stoehrmann, & Radandt, 1998).

The general potential of mnemonic strategies can be extended to benefit a wider range of students, including students with intellectual and developmental disabilities, in several ways. Research has found that students with intellectual disabilities “show increased learning and memory when the content is presented in meaningful contexts” (Taylor & Turnure, 1979, p. 660). Mnemonic strategies provide a means for students to overlay context meaningful to them in situations that might otherwise not be the case. Mnemonic strategies can tap areas of potential cognitive strength (memory for pictures, acoustic memory) for students with intellectual impairments, while de-emphasizing relative weakness (Scruggs, Mastropieri, & Levin, 1987) and enable students to learn new skills or information in a way that is more meaningful to them through acoustic-imaginal linking (Mastropieri, Scruggs, & Levin, 1985). In addition, mnemonic strategies have been used for addressing problem behavior, which is often a barrier to access to the general curriculum for students with intellectual and developmental disabilities (Carpenter, 2001). Mnemonic strategies are used in this manner in two ways: (a) by being infused into student-directed learning strategies, such as problem solving, self-instruction, and self-monitoring, through keyword or letter strategy mnemonics based on imagery or acoustic linking strategies (Silverstein, 1997; Smith et al., 1994), and (b) by being infused into positive behavior support models by arranging antecedent events of the behavior problem or by providing appropriate support to replace the function of the behavior problem through a visual card or social mnemonic strategy (Agran et al., 2003).

Another potential for mnemonic strategies is that they can be applied to different levels of curriculum modifications. For example, keyword and letter methods (e.g., using acronyms) through visual images or sketches can be applied as a curriculum adaptation (Bulgren & Lenz, 1996) by being incorporated into graphic organizers. Moreover, students can be taught to generate their own mnemonic strategies, including keyword strategies, rhymes, or acoustic linking strategies that, in turn, involve a curriculum augmentation strategy (Wehmeyer et al., 2001). Mnemonics also contribute to the efficacy of other curriculum augmentation strategies, such as student-directed learning strategies (Agran et al., 2003; Smith et al., 1994; Wehmeyer, Sands et al., 2002).

Student-Directed Learning Strategies and Self-Determination

Student-directed learning strategies form a subset of broader learning or cognitive strategies, and represent a powerful means to augment the curriculum to enable students with intellectual and developmental disabilities to perform more effectively in the general curriculum (Wehmeyer et al., 2001). Moreover, promoting and enhancing self-determination and its component elements (goal-setting, problem-solving, self-regulation and other skills) equips students with disabilities with skills that, in turn, will enable them to succeed in the general curriculum (Wehmeyer et al., 2004). Student-directed learning strategies
(Agran et al., 2003) enable students to learn to direct their learning and self-regulate problem solving geared toward learning (Wehmeyer, Palmer, Agran, Mithaug, & Martin, 2000). Teaching students strategies such as antecedent cue regulation, self-instruction, self-monitoring, self-evaluation, and self-reinforcement has multiple benefits, including promoting inclusion and self-regulated learning (Agran et al., 2003). There is now a fairly robust body of literature documenting the impact of promoting self-determination and student-directed learning on positive outcomes for children and youth with intellectual and developmental disabilities (Agran et al., 2003; Algozzine, Browder, Karvonen, Test, & Wood, 2001; Wehmeyer, Abery, Mithaug, & Stancilfie, 2003), and an emerging database suggesting that such strategies result in enhanced access to the general curriculum (Palmer et al., 2004; Wehmeyer et al., 2004).

**Goal-Setting**

A critical component of self-determined behavior is goal-setting and attainment. Goal-setting involves: a) identifying and defining a goal, b) developing an action plan that consists of specific steps that will be undertaken in an effort to achieve the goal, and c) evaluating the outcomes of these actions (Locke & Latham, 1984, 1994). A goal is, in essence, a specification of what a person wishes to achieve through his or her actions. Goals act to regulate our actions (Locke & Latham, 2002). As an augmentation to the general curriculum, teaching students with intellectual and developmental disabilities to set and attain goals can enable them to better regulate their behavior as it relates to their academic progress by providing them with an established criterion in which to compare their current level of performance.

Research on the impact of goal-setting on academic performance has focused primarily on students with learning disabilities. This literature base indicates that goal-setting interventions have a positive impact on the academic performance of students across a variety of academic domains, including writing (Graham, MacArthur, Schwartz, & Page-Voth, 1992; Page-Voth & Graham, 1999; Troia & Graham, 2002), arithmetic (Schunt, 1985), and spelling and vocabulary evaluation (Gardner & Gardner, 1978). The goal-setting interventions that have been evaluated range from very simple goal-setting interventions, where a teacher or researcher simply asked students to state a performance goal prior to beginning an assignment or studying for a test (Gardner & Gardner; Schunt), to specific, structured strategies, such as Do PLANS (Pick goals, List ways to meet goals, And, make Notes, Sequence notes) (Graham et al.), STOP & LIST (Troia & Graham), or the SMG (Student Management Guide) (Lenz, Ehren, & Smiley, 1991).

Even though there is limited research about the impact of goal-setting on the academic performance of students with intellectual and developmental disabilities, goal-setting has promise to promote greater access to the general curriculum for this population. There is limited, though emerging, evidence that students with intellectual disabilities can be taught the skills necessary to set and achieve goals (Copeland, Hughes, Agran, Wehmeyer, & Fowler, 2002; German, Martin, Marshall, & Sale, 2000), although most of these evaluations have been with non-academic content areas (Copeland & Hughes, 2002). There is ample evidence, though, that students with intellectual and developmental disabilities can be taught the skills necessary to set and achieve goals (Mithaug, Mithaug, Agran, Martin, & Wehmeyer, 2003; Wehmeyer, Abery, et al., 2003; Wehmeyer et al., 2000). As noted earlier, Palmer et al. (2004) taught middle school students with intellectual disabilities to set and attain goals linked to grade-referenced standards in the general curriculum, illustrating the potential role that promoting goal setting has in achieving access to the general curriculum for this population. Like other strategies, goal-setting also can be incorporated into other adaptation or augmentation strategies.

**Problem-Solving**

A second component element of self-determined behavior (Wehmeyer, 2001), problem-solving, also has promise to promote student access and progress in the general curriculum. Problem-solving is a process used to identify available information and design solutions to
a problem in order to achieve one’s goal (Agran, Blanchard, Wehmeyer, & Hughes, 2002). Generally, four steps are involved in a traditional problem solving process: (a) identify the problem, (b) identify potential solutions to the problem, (c) identify barriers to solving the problem, and (d) identify consequences to each solution (Agran & Wehmeyer, 1999). Teaching problem solving is a critical element not only for ensuring students’ success in general education (Agran, Blanchard et al., 2002), but also for school reform efforts (Gumpel, Tappe, & Araki, 2000; Peterson, 1996).

Teaching problem-solving skills has been used as a curriculum augmentation for students with learning disabilities. Such instruction forms the core of many strategic instruction activities validated with students with learning disabilities, focusing on the problem-solving aspects of using information or knowledge (Deshler, Ellis, & Lenz, 1996). Problem solving is an especially important skill for math (Gersten & Baker, 1998; Maccini & Hughes, 2000). Gersten and Baker demonstrated that problem-solving strategies were also useful for students with learning disabilities in learning science content.

Similar to goal-setting, there is evidence, albeit limited, that students with intellectual and developmental disabilities can learn to solve problems (Wehmeyer, Agran, Palmer, Mithaug, & Martin, 2003), including in the context of the general education classroom (Agran, Blanchard et al., 2002). While most of the research on problem solving with students with intellectual and developmental disabilities is related to non-academic content such as transition skills, workplace behaviors, and community and leisure activities (Agran, Blanchard, & Wehmeyer, 2000; Hughes & Rusch, 1989; O’Reilly, Lancioni, & Kierans, 2000), Palmer et al. (2004) showed that instruction in problem solving could enable students with intellectual and developmental disabilities to make progress in the general curriculum.

Again, integrating problem-solving into other augmentation strategies is an effective means to achieve more positive outcomes. For example, traditional methods of teaching self-instruction essentially teach students to articulate a problem-solving sequence (Hughes, Hugo, & Blatt, 1996). In addition, problem-solving can be incorporated into goal-setting as shown by research with the Self-Determined Learning Model of Instruction (Wehmeyer, Abery et al., 2003; Wehmeyer et al., 2000). Also, problem-solving can be used with technology. Mastropieri, Scruggs, and Shian (1997) demonstrated that students with mild mental retardation can learn mathematical problem-solving skills through a computer animation program. The students learned problem solving more effectively in computer-based learning modes than in paper-pencil based learning activities.

Table 1 provides a brief summary of the potential of each aforementioned strategy for students with intellectual and developmental disabilities and practical suggestions to apply the strategies in the classroom.

Conclusions

We would suggest that there is a need, given the impetus provided to the field by the IDEA Access to the General Curriculum mandates, to engage in research and model development to examine the efficacy of the strategies described above to support students with intellectual and developmental disabilities to become involved in and progress in the general curriculum, and to develop models that modify these strategies or create new strategies. As a first step, current studies on the effect of strategies for access for students with disabilities need to “shift gears” to some degree to focus research and model development that moves away from examining achievement only in non-academic content areas and moves toward examining outcomes in academic content areas, as well as move away from models that rely exclusively on external supports towards models that focus on the incorporation of curriculum augmentation strategies that enable students to more effectively teach themselves. In so doing, we believe, students with intellectual and developmental disabilities will be more effective in achieving access to and progress in the general curriculum.
<table>
<thead>
<tr>
<th>Strategies</th>
<th>Potential/findings(*) for students with intellectual and developmental disabilities</th>
<th>Suggestions for applications in classroom</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graphic organizer</strong></td>
<td>● Can be used as a curriculum adaptation and augmentation across various content areas.</td>
<td>● Use written advance organizer to introduce purpose for a big ideas or class lesson.</td>
</tr>
<tr>
<td>Types: Flow chart, Semantic maps, Webs, Computerized program</td>
<td>● Written organizer is more effective than oral organizer.* ● Concisely introducing a purpose statement helps students’ comprehension.* ● Can be used to facilitate student participation through various group formats.</td>
<td>● Provide pictures or illustrations as a graphic organizer for students who are not good at reading. ● Use graphic organizers as handouts for the class or homework that have students fill in blanks that are included in the graphic organizer. ● Simplify graphic organizers by reducing complexity, clarifying symbols, and providing enough space. ● Use graphic organizer as a small group activities to maximize student participation and interaction with other peers. ● Use multimedia technology to extend the potential of graphic organizers. ● Use chunking with graphic organizers to combine related elements into one unit. ● For students who are not good at reading, use pictures and other visual cues. ● Provide opportunities to chunk the student’s favorite activities or goal related to the IEP by using pictures or photo. ● Incorporate chunking with other self-directed learning strategies. ● Gradually fade teacher-directed teaching to empower students.</td>
</tr>
<tr>
<td><strong>Chunking</strong></td>
<td>● Can be used as a curriculum adaptation and augmentation across various content areas.</td>
<td>● Can be used for enhancing attention and memory for students.</td>
</tr>
<tr>
<td></td>
<td>● Can be incorporated with other strategies, such as graphic organizers and self-management skills.</td>
<td>● Should be taught directly, but gradually fading the teacher’s intervention to enable students to be more active learners</td>
</tr>
<tr>
<td></td>
<td>● Can be more effective when incorporated with visual cues and technology.</td>
<td>● Identify familiar part from new information and provide visual cues related to them to provide new information in a meaningful way</td>
</tr>
<tr>
<td><strong>Mnemonic strategies</strong></td>
<td>● Can meet individual needs based on characteristics of students with intellectual and developmental disabilities, including intelligence and adaptive behavior.</td>
<td>● Use antecedent cue regulation strategies, such as picture or auditory prompts to enable students to manage behavior in class.</td>
</tr>
<tr>
<td>Types: Imagery devices, Word-based devices (Keyword, Pegword, Letter method) Combining mnemonic devices</td>
<td>○ Are effective in learning and memorizing new information through memory for picture or acoustic memory that is cognitive*</td>
<td>● Provide handout or worksheet including highlighted or colored keyword.</td>
</tr>
<tr>
<td>Strategies</td>
<td>Potential/findings(*) for students with intellectual and developmental disabilities</td>
<td>Suggestions for applications in classroom</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
</tbody>
</table>
| ○ Are effective in managing behavior including self-regulatory management and PBS* | * Incorporate mnemonic strategies to self-instruction (self-talk), problem-solving, and goal-setting.  
   ● Can be applied to different levels of curriculum modification, adaptation, augmentation and alternation in non-academic context and in academic context* | * Modify/apply mnemonic strategies widely used below:  
   ○ Keyword method  
     ● As a starting point, enable students to recognize and remember keyword itself instead of new words that are related to the keyword  
     ● Use keyword that is related to students’ meaningful experiences or familiar environments.  
     ● Incorporate keyword method to identify goal or problem  
     ● Present or highlight the keyword with verbal cues and visual images, such as a card or photo, etc  
   ○ Pegword method  
     ● Incorporate pegword method to keyword method, as possible, instead of using it alone  
     ● Infuse pegword into a familiar song or melody  
   ○ Letter method  
     ● Create a short question with a picture or visual card considering students’ learning styles, needs and abilities with visual images, such as a card or photo, etc  
     ● Teach self-instruction and problem-solving with a letter mnemonic to be practiced in meaningful classroom activities  
     ● Present letter with visual cards and verbal cues |
| ● Can be used in academic context and can be incorporated within problem-solving strategies. | * Can be used as a curriculum augmentation.  
   ● Can be incorporated within problem-solving strategies.  
   ● Students with intellectual disabilities can learn the skills necessary to set and achieve goals even though such skills are improved performance in non-academic domains (e.g., vocational and behavioral)* | * Provide clear purpose or objectives for class activities across various content areas  
   ● Enable students to set a personal goals based on their preference and interests to motivate learning and increase engagement.  
   ● Provide options or choices for students who are not good at expressing their own preference by using visual cues  
   ● Enable students to participate in IEP meeting to understand their future goal in general classroom  
   ● Use visual cues, graphic organizers or chunking to clarify priorities of target goal  
   ● Provide the opportunity to set a goal for a class, with peers in small groups |
| ● Can be used to motivate students by enabling them to set their own goal according to their preference. | * Can be used with visual structuring system. | * Select visual structuring system that is meaningful to students with intellectual disabilities and who have a history of disliking certain familiar objects or symbols.  
   ● Enable students to participate in IEP meeting to understand their future goal in general classroom  
   ● Use visual cues, graphic organizers or chunking to clarify priorities of target goal  
   ● Provide the opportunity to set a goal for a class, with peers in small groups |
### References


### TABLE 1—(Continued)

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Potential/findings(*) for students with intellectual and developmental disabilities</th>
<th>Suggestions for applications in classroom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem-solving</td>
<td>● Can be used as a curriculum augmentation. ● Can be incorporated into goal setting. ● Can be generalized across different class activities. ● Enables students to reach their own goals, especially related to non-academic content such as transition and workplace skills, community and leisure activity.* ● Is effective for students to achieve their goals in academic areas, such as following direction and class contribution.* ● When combined with other strategies, such as self-instruction, is more effective.*</td>
<td>● Help students identify their own problems in academic areas by providing them with an opportunity to choose problems that are related to IEP goals. ● Use pictures or other visual cues to help students express their opinions about problem or goal ● Use a systematic learning program related to problem-solving such as the Self-Determined Learning Model of Instruction. ● Incorporate self-instruction, self-monitoring and mnemonic strategies for effective learning of problem-solving. ● Teach problem-solving skills in the natural context that students may face in regular routines. ● Provide students with opportunities to choose a reward for solving the problem. ● Enable students to solve problem with other peers in small group.</td>
</tr>
</tbody>
</table>

208 / Education and Training in Developmental Disabilities-September 2006


Guastello, E. F., Beasley, T. M., & Sinatra, R. C. (2000). Concept mapping effects on science content comprehension of low-achieving inner-city...


acquisition: Teaching generalization for students with developmental disabilities. In A. Hilton & R. Ringlaben (Eds.), Best and promising practices in developmental disabilities (pp. 139–155). Austin, TX: PRO-ED.


Received: 4 April 2005
Initial Acceptance: 25 May 2005
Final Acceptance: 30 August 2005
Functional Communication Training: A Review of the Literature Related to Children with Autism

G. Richmond Mancil
University of Florida

Abstract: Numerous researchers have employed functional communication training (FCT) to address both the communication and behavioral needs of children with autism. Thus, the purpose of this review is to examine FCT, particularly, the environments and individuals involved in the training and the effectiveness of FCT with children who have a diagnosis of Autism Spectrum Disorder (ASD) and to provide suggestions for practitioners and researchers. FCT consistently reduces challenging behavior and increases communication; however, the majority of research is clinically based and focuses on one communication mand. Future research teams should address maintenance and generalization by training teachers in classrooms and parents in homes while collecting data across time.

Autism is a developmental disorder affecting the lives of thousands of children. According to the Centers for Disease Control and Prevention (CDC), approximately 34 in 10,000 children ages 3 to 10 years of age have autism (CDC, 2004). The Autism Society of America (ASA) reports that 1 in 166 babies born today will develop autism. The ASA also notes that 1.5 million Americans including children and adults have autism, while another 15 million Americans (e.g., family, educators, and health care workers) are affected by autism.

The essential features of autism include a significant impairment in social interaction and communication and a highly restricted area of activities and interests (American Psychiatric Association, 2000). Concurrent with these features, children with autism exhibit high levels of challenging behaviors such as screaming, hitting, and biting (Sigafoos, 2000), thus, creating substantial obstacles for individuals charged with their education and well being (Durand & Merges, 2001). For example, many parents experience stress when their children engage in tantrums. Unlike other parents, parents of children with autism usually cannot determine the reason for the tantrum because of their child’s deficits in communication. Such issues with communication deficits and challenging behaviors combined with the increase in the prevalence of autism demands for the field of behavior disorders to respond and provide evidence-based practices to meet these children’s needs at home and in educational settings.

Several researchers have responded by looking at challenging behaviors and their relationship with communication abilities (e.g., Bott, Farmer, & Rhode, 1997; Chung, Jenner, Chamberlain, & Corbett, 1995; Sigafoos, 2000; Schroeder, Schroeder, Smith, & Dalldorf, 1978). Chung et al. found an inverse relation between communication ability and the display of challenging behaviors such as self-injury and aggression. Similarly, Bott et al. (1997) discovered that individuals with more developed speech skills had a lower frequency of challenging behaviors than those with impaired speech. Further, Sigafoos hypothesized in a more recent study that impaired communication development causes challenging behaviors.

To address both the communication and behavioral needs of children with autism, several researchers employed functional communication training (FCT) (Carr & Durand, 1985; Durand & Merges, 2001; Wacker et al., 1990). Developed in the mid-1980s, FCT in-
involves assessing the function of a behavior (e.g., attention, escape, tangible, or sensory) through functional behavior assessments (FBA) and then replacing the challenging behavior with a communicative response that serves the same function.

FBA typically consists of interviews, direct observations, and functional analyses (Brady & Halle, 1997). Interviews involve asking teachers, parents, and other caregivers communicative and behavioral-related questions. The next step consists of directly observing the behavioral and communicative behaviors of the children in various natural settings such as the classroom and playground. Both of these steps aid in developing a hypothesis of the behavior’s function. Finally, the trainer completes a functional analysis (FA) by manipulating variables such as demands, attention, and tangible items to see if performing the challenging behavior allows the child to escape the difficult task or gain attention.

After completing the FBA process, trainers must identify a communicative response. This communicative response may consist of a response from one of the following categories: verbal language, picture communication, gestures, or assistive technology devices (Brady & Halle, 1997). The selection of this response should be based on the child’s capability of completing the response, the ease of teaching the response, the acknowledgement from others of the response, and how quick the response serves its function (Dunlap & Duda, 2005; Horner & Day, 1991).

After the response selection, the actual communicative response should be taught (Lalli, Casey, & Kates, 1995). For example, a child may be taught to say, help for assistance (i.e., obtaining attention) instead of screaming for help. Also, a child may be taught to give a picture of a requested item (i.e., obtain a tangible) to the teacher for access to the item instead of screaming or hitting to gain the tangible. In both cases, the instructor must ensure mastery of the response before proceeding further.

The final step in FCT involves ignoring the challenging behavior (i.e., the use of extinction) and prompting and acknowledging the use of the communicative response that replaces the challenging behavior (Lalli et al., 1995). For example, while ignoring a tantrum a teacher may prompt a child to ask for a break and then provide the break after the child responds with the appropriate communicative response. Researchers purport that this process increases communication and decreases challenging behaviors (Carr & Durand, 1985; Durand & Merges, 2001; Wacker et al., 1990). The majority of the FCT research, however, has been conducted with children with severe or profound disabilities (e.g., severe, profound mental retardation), not individuals solely identified as having autism, thus, making it difficult to generalize the findings to children with autism. To complicate matters, there is little research on training parents and teachers how to use FCT and most of the research is from clinical type settings, not less structured settings like typical classrooms and homes. Thus, the purpose of this review is to examine functional communication training, particularly, the environments and individuals involved in the training and the effectiveness of FCT with children who have a diagnosis of Autism Spectrum Disorder (ASD). First, analyses of the characteristics of the participants, environment, research designs, behaviors, interventions, major findings, reliability, and treatment fidelity across studies are provided. Second, a critique of the findings to address limitations and implications for future researchers are provided.

Method

The literature review consisted of searches of ERIC, Education, PsycINFO, and Academic Search Premier data bases using various combinations of the following terms: functional communication training, functional equivalence training, autism, autism spectrum disorder, and communication. Then a hand search was conducted of the following journals, covering the span of 1985 to the present: Focus on Autism and Other Developmental Disabilities, Journal of Autism and Developmental Disorders, Journal of Applied Behavior Analysis, and Journal of Positive Behavior Interventions. Finally, after reviewing the retrieved articles, an archival search was conducted. These searches produced 30 articles in which FCT was the primary intervention addressing challenging behaviors and communication deficits. The
following criteria were used for inclusion in this review: (a) at least one participant of the study was a child with an autism spectrum disorder diagnosis, (b) the function of the challenging behavior was determined by the functional behavior assessment (FBA) process, and (c) the primary intervention was functional communication training.

Although some studies included children and adults with varying disabilities, only the children diagnosed with ASD were included in the analyses. From the list of 30 articles initially identified, eight studies were identified in which all participants had received prior diagnoses of autism. For each study, characteristics of the participants, research environment, research designs, behaviors, interventions, major findings, reliability, and treatment fidelity were identified. These data are presented in Tables 1, 2, 3, and 4.

Results

Characteristics of Study Participants

The majority of researchers reported gender along with chronological, language, and mental ages. Eighteen of the participants were male and four were female. Although this may at first seem overrepresented by males, it represents the overall population of individuals diagnosed with autism (Autism/Pervasive Developmental Disorders’ Newsletter, 2003). As shown in Table 1, participants ranged in age from 2.7 to 13 years, with the average age being 8 years. The language age reported ranged from 1.8 to 13.3 years, while the mental age reported ranged from 2.4 to 7.9 years. Based on this data, it can be inferred that no relationship exits between chronological, mental, and language age when compared across participants. For example, a participant with a low chronological age could have a language and mental age higher than that of an older participant. Also, while some participants had a mental age greater than their language age, other participants’ ages were opposite in correlation.

Participants differed in their levels of language prior to the studies and whether or not they had concomitant diagnoses. Of the 12 prior speech levels reported, five participants (42%) spoke in complete sentences; however, their sentences were not always functional. For example, some of the complete sentences were echolalic (i.e., repeated phrase over and over), while others were bizarre (e.g., The cat flew on a broom.). Also, of the individuals who spoke in complete sentences, only 1 (10%) was reported to speak spontaneously. Two participants (16%) were nonverbal and did not communicate with signs or gestures. Of all these participants, only six (27%) were reported to have additional diagnoses, which included communication disorders, seizure disorders, and severe/profound mental retardation.

Characteristics of Environment

With a few exceptions and unlike participant characteristics, the environmental context of the studies was similar. As shown in Table 2, researchers and research assistants implemented the majority of studies. Only one study (12.5%) involved a teacher as an implementer, and she was not the participant’s teacher, that is, she acted more as a research assistant (Wacker et al., 1990). Further, only one study (12.5%) used parents as trainers (Wacker et al., 2005). Similarly, training occurred in clinic rooms. For example, the majority of training (i.e., 6 studies, which is 75%) occurred in separate rooms that usually contained only a table and chairs. Two studies (25%), however, were conducted in more natural environments. For example, O’Neill and Sweetland-Baker (2001) conducted their study in various locations of the participant’s classroom while Wacker et al. conducted training in designated rooms of the participant’s home.

Characteristics of Research Designs, Behaviors, and Interventions

Challenging behaviors and their functions varied among participants. Consequently, researchers implemented various interventions to address these functions. Specifically, differences identified across the studies can be organized in the following categories: (a) challenging behaviors, (b) FBA procedures and designs, (c) behavior functions, (d) communicative responses, and (e) FCT research designs.
### TABLE 1

**Characteristics of the Study Participants**

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>CA*</th>
<th>LA*</th>
<th>MA*</th>
<th>M</th>
<th>F</th>
<th>Diagnosis</th>
<th>Prior Speech Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carr &amp; Durand (1985)</td>
<td>1</td>
<td>13</td>
<td>N/A</td>
<td>3</td>
<td>1</td>
<td></td>
<td>Autism</td>
<td>Verbal (complete sentences)</td>
</tr>
<tr>
<td>Durand &amp; Carr (1992)</td>
<td>3</td>
<td>3.8–4.9</td>
<td>1.8–3.8</td>
<td>2.4–4.4</td>
<td>2</td>
<td>1</td>
<td>Autism</td>
<td>N/A</td>
</tr>
<tr>
<td>Martin et al. (2005)</td>
<td>1</td>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td></td>
<td>Autism</td>
<td>Nonverbal</td>
</tr>
<tr>
<td>O’Neill &amp; Sweetland-Baker (2001)</td>
<td>2</td>
<td>6–15</td>
<td>N/A</td>
<td>N/A</td>
<td>2</td>
<td></td>
<td>Autism</td>
<td>Verbal (non-functional)</td>
</tr>
<tr>
<td>Wacker et al. (1990)</td>
<td>1</td>
<td>7</td>
<td>N/A</td>
<td>N/A</td>
<td>1</td>
<td></td>
<td>Autism</td>
<td>None</td>
</tr>
<tr>
<td>Wacker et al. (2005)</td>
<td>7</td>
<td>2.7–6.5</td>
<td>N/A</td>
<td>N/A</td>
<td>6</td>
<td>1</td>
<td>3- Autism; 4- PDD</td>
<td>N/A</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>(2.7–15)</td>
<td>(1.8–13.3)</td>
<td>(4.1–7.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** CA = chronological age; LA = language age; MA = mental age; N/A = not available; MR = mental retardation; func. = functional; M = male; F = female; m = mean

* age in years/months
Challenging behaviors. Teachers and parents identified 8 different categories of challenging behaviors, with many participants exhibiting more than one behavior. The categories included aggression (e.g., hitting, hair pulling), self-injurious behavior (e.g., hand biting), destroying property, tantrums (e.g., yelling), body rocking, hand flapping, oppositional behavior (e.g., refuse to do work), and walking away. Fourteen participants (64%) exhibited aggression, self-injurious behavior, or destruction of property. Also, fourteen participants (64%) exhibited more than one challenging behavior. For example, one participant engaged in aggressive behavior, tantrums, self-injurious behavior, and property destruction (Carr & Durand, 1985).

FBA procedures. Basic FBA procedures used to analyze various behaviors were fairly similar across all studies. First, information was collected through interviews with teachers or parents, however interviews used in various studies differed in length. While some researchers reported interviews that were pages long, others reported interviews as short as one to two questions. Next, direct observations were conducted to further aid in developing a hypothesis of the behavior’s function. Finally, a functional analysis (FA) was conducted to determine the function. However, designs used for the FA differed. As shown in Table 3, five research teams used alternating treatment designs where the assessment conditions were systematically alternated. For example, Carr and Durand (1985) alternated easy versus difficult tasks and a low (33%) versus high (100%) attention condition. Similarly, Wacker et al. (1990) alternated escape, tangible, alone, and social attention conditions. Durand and Carr (1987; 1992) used a reversal design to examine effects of different conditions. Wacker et al. (2005) used a multi-element design to compare assessment conditions (i.e., attention, escape, tangible, and free play) by counterbalancing them across sessions.

Behavioral functions. After the completion of the FBA procedures, behavioral functions were definitively identified in each study for all but one participant (Wacker et al., 2005) whose behavioral function was determined to be undifferentiated. Across all the studies, 12 participants (55%) emitted challenging behaviors to escape a task or situation, while eight participants (36%) displayed challenging behaviors to gain attention. Only three participants (14%) engaged in challenging behaviors to gain a tangible. Also, of the 12 participants whose behavioral function was escape, four of them (18%) also engaged in the behavior to gain attention. Similarly, one participant (5%) who engaged in challenging behaviors did so to gain attention and to gain access to a tangible object.

### Table 2

**Characteristics of the Environment**

<table>
<thead>
<tr>
<th>Study</th>
<th>FCT Implementer</th>
<th>FCT Training Location</th>
<th>Description of Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carr &amp; Durand (1985)</td>
<td>Researcher</td>
<td>Separate Class</td>
<td>5 × 10-m classroom next door</td>
</tr>
<tr>
<td>Durand &amp; Carr (1987)</td>
<td>Researcher/ Assistants</td>
<td>Separate Class</td>
<td>Contained table, two chairs</td>
</tr>
<tr>
<td>Durand &amp; Carr (1992)</td>
<td>Research Assistants</td>
<td>Separate Class</td>
<td>Table, chairs for trainer, children</td>
</tr>
<tr>
<td>Martin et al. (2005)</td>
<td>Researcher</td>
<td>Separate Class</td>
<td>N/A</td>
</tr>
<tr>
<td>O’Neill &amp; Sweetland-Baker (2001)</td>
<td>Researcher</td>
<td>Various class locations</td>
<td>N/A</td>
</tr>
<tr>
<td>Ross (2002)</td>
<td>Researcher</td>
<td>Separate Class</td>
<td>Table, bookshelf, 3 chairs</td>
</tr>
<tr>
<td>Wacker et al. (1990)</td>
<td>Therapist, graduate students, 1 teacher</td>
<td>Standard classroom/therapy room</td>
<td>Therapy room had table, chairs</td>
</tr>
<tr>
<td>Wacker et al. (2005)</td>
<td>Children’s parents</td>
<td>Room at home</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*Note.* N/A = not available.
### TABLE 3

Research Designs and Interventions

<table>
<thead>
<tr>
<th>Study</th>
<th>FCT Design</th>
<th>FBA Procedures</th>
<th>Dependent Measures</th>
<th>Function</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin et al. (2005)</td>
<td>Reversal</td>
<td>FA Alt. Tx</td>
<td>TAN, AG, W</td>
<td>E</td>
<td>Picture Card</td>
</tr>
<tr>
<td>Ross (2002)</td>
<td>Reversal</td>
<td>FA Alt. Tx</td>
<td>no or poor initiation</td>
<td>1-A, 1-E, 1-T</td>
<td>Verbal</td>
</tr>
<tr>
<td>Wacker et al. (1990)</td>
<td>Reversal</td>
<td>FA Alt. Tx</td>
<td>SIB</td>
<td>T</td>
<td>Sign language</td>
</tr>
<tr>
<td>Wacker et al. (2005)</td>
<td>Multiple Baseline</td>
<td>FA multiple</td>
<td>6-AG; 4-SIB; 3-DP</td>
<td>4-A &amp; E; 1-A &amp; T; 1-U</td>
<td>Verbal, Assistive Technology</td>
</tr>
</tbody>
</table>

*Note.* AG = aggressive behavior; SIB = self-injurious behavior; DP = destroying property; OP = oppositional; TAN = tantrum; W = walk away; FCT = functional communication training; FA = functional analysis; FBA = functional behavioral assessment; HF = hand flapping; BR = body rocking; N/A = not available; I = interview; O = observation; Alt. Tx = alternating treatment; A = attention; E = escape; T = tangible; U = undifferentiated
<table>
<thead>
<tr>
<th>Study</th>
<th>Tx Fidelity</th>
<th>Inter Rater Reliability</th>
<th>Behavioral Results</th>
<th>Communication Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carr &amp; Durand (1985)</td>
<td>High</td>
<td>80% or higher</td>
<td>DB decreased to 0.5%</td>
<td>Relevant responses maintained</td>
</tr>
<tr>
<td>Durand &amp; Carr (1987)</td>
<td>High</td>
<td>80% or higher</td>
<td>Reduction in HF, BR exhibited following training</td>
<td>Communication increased</td>
</tr>
<tr>
<td>Durand &amp; Carr (1992)</td>
<td>High</td>
<td>N/A</td>
<td>Decreased, maintained best in FCT/time-out</td>
<td>Unprompted communication</td>
</tr>
<tr>
<td>Martin et al. (2005)</td>
<td>High</td>
<td>97–99%</td>
<td>Bear hugging decreased</td>
<td>Independently after A phase</td>
</tr>
<tr>
<td>Ross (2002)</td>
<td>High</td>
<td>88–100%</td>
<td>N/A</td>
<td>Faulty responses decreased</td>
</tr>
<tr>
<td>Wacker et al. (1990)</td>
<td>High</td>
<td>92% average</td>
<td>FCT w/time-out resulted in hand biting decreasing to 0%</td>
<td>Signing was maintained</td>
</tr>
<tr>
<td>Wacker et al. (2005)</td>
<td>High</td>
<td>90–100%</td>
<td>Behaviors decreased for all participants ranging from 66.25% to 100% reduction</td>
<td>Manding increased</td>
</tr>
</tbody>
</table>

*Note.* Tx = treatment; HF = hand flapping; BR = body rocking; N/A = not available; FCT = functional communication training; Db = disruptive behavior.
Communication responses. With the knowledge of the behavioral function(s), trainers taught an array of communicative responses to replace the challenging behaviors, with equal success regardless of the response category. Responses taught aligned with the identified function and fit into one of the following categories: (a) verbal language, (b) sign language, (c) picture icon based language, or (d) augmentative devices. As shown in Table 3, trainers in six studies taught participants to verbally mand. For example, Durand and Carr (1987) taught participants to verbally mand, *Help me* to replace body rocking and hand flapping that served to escape aversive tasks and situations. Durand and Carr (1992) taught participants in another study to verbally mand, *Am I doing good work?* to replace tantrums that served to gain attention. Two research teams taught students to use sign language for manding (Wacker et al., 2005; Wacker et al., 1990). For example, to gain access to a tangible, Wacker and colleagues (1990) taught participants to mand the sign *please*. In two studies, research teams taught students to use icons to serve as the mand (Martin, Drasgow, Halle, & Brucker, 2005; Wacker et al., 2005). For example, Martin et al. taught their participant to present an icon card with the words *No Thank You* written upon it to replace tantrums, which served as an escape function. Wacker et al. (2005) used an augmentative device to teach a participant to press a micro switch that said, *Please*. This response replaced self-injurious behavior, which was attention maintained.

Across all studies, communication responses were taught in a similar fashion. For example, when Durand and Carr (1992) taught participants to verbally mand, *Am I doing good work?* the training continued until the children were able to perform the task correctly 10 consecutive times. Similarly, in the study conducted by Martin and colleagues (2005), participants were trained to emit the communicative response without error. In summary, all participants were taught one mand until they were able to respond correctly 10 consecutive times.

FCT designs and procedures. After the communicative responses were taught, researchers used one of the following research designs for FCT: multiple baseline or reversal. As shown in Table 3, researchers used a reversal design in four studies. For example, Carr and Durand (1985) alternated relevant and irrelevant response phases with baseline. First, baseline data were collected in which no intervention was implemented. Then, participants were reinforced for relevant communicative responses by the trainer giving them assistance. After this phase, reinforcement was removed, thus, returning to baseline. Next, participants were reinforced for irrelevant responses; these responses however had nothing to do with the task at hand or the behavior’s function. The sequence was then repeated; however, the phases (i.e., relevant and irrelevant) were counterbalanced. Also as shown in Table 3, four studies used multiple baseline design. For example, Durand and Carr (1992) implemented the procedures with one participant as they continued to collect baseline data on the remaining participants. After a few sessions, they implemented the procedures with the next participant as they continued to implement the procedure with the first participant. This continued until the procedure was implemented with every participant.

Major Findings

Regardless of the research design, research teams reported similar findings, in that they were all successful regardless of the topography of the behavior (i.e., aggression, tantrums, self-injurious behavior), the behavior’s function, or the mode of communication. As shown in Table 4, this success was demonstrated by a decrease of challenging behaviors with a corresponding increase in communication, but the communication mands were limited in scope. For example, in a study by Carr and Durand (1985), a participant’s challenging behaviors decreased and his use of one mand increased.

Behavioral results. After implementation of FCT, research teams found a decrease in challenging behavior across all studies. For example, Wacker and colleagues (2005) found that FCT combined with time-out resulted in hand biting decreasing to zero percent. Similarly,
Carr and Durand (1985) found challenging behaviors to decrease to 0.5% upon the successful implementation of FCT. Also, Wacker and colleagues found significant decreases in challenging behavior for all but one participant, whose behavior’s function was undifferentiated, consequently, researchers noted that the communication response must match the challenging behavior’s function.

**Communication results.** Akin to the positive behavioral results across studies, participants increased their use of communication mands, but they were limited in range. Carr and Durand (1985) noted sustained rates of relevant responses in the final phases of the study, but only one response was taught and measured. In a later study, Durand and Carr (1992) also found an increase in unprompted communication, the communication however was again limited to one type of mand. More recently, the participant in Martin et al.’s (2005) study independently used his icon card to request an item 100% of the time following training in phase; similar to previous studies, they focused on one mand with no expansion of communication.

**Reported Reliability and Treatment Fidelity**

Reported research results would be compromised unless the procedures were implemented with fidelity and observations were reliable. Fortunately, all research teams reported high treatment fidelity (i.e., consistence of implementation of the methods and procedures of treatment) within their prospective studies. They also reported high inter-rater reliability (i.e., agreement of observed phenomenon). For example, Carr and Durand (1985) reported reliability of 80% or greater for all categories. Similarly, Wacker et al. (1990) had an average reliability of 92%, with 80% or more for each category. Martin and colleagues (2005) reported greater reliability ranging from 97% to 99%.

**Discussion**

A significant impairment in communication is one of the defining characteristics of autism, subsequently causing problems with behavior. FCT is one approach researchers employed to address the communication and behavioral needs of children with autism. As researchers developed FCT, they provided interventions in clinical settings removed from natural environments (e.g., children’s classrooms, homes), which is typical for the initial stages of procedural development. When implementing the procedures during initial development stages, research teams produced positive behavioral and communication results. For example, Durand and Carr (1987) indicated an increase in communication and a decrease in challenging behaviors, but this study occurred within a separate 5 x 10 meter classroom that excluded the child’s teacher and focused on one communication mand. Thus, readers may acknowledge that FCT works when researchers conduct training in small isolated rooms and focus on one communication mand.

Similarly, other research teams indicated an increase in communication and a decrease in challenging behaviors (e.g., Carr & Durand, 1985; Durand & Carr, 1987; 1992). Based on the evidence provided in this review with the progression of knowledge and time, researchers for the other studies continue to conduct FCT similar to the first published article in 1985. They typically conducted the training themselves and did not extend the research beyond clinical settings that focus on one communication mand. Each limitation poses a critical threat to maintenance and generalization of the communication and behavioral results, consequently decreasing the effectiveness of the FCT.

**Maintenance**

Although children achieve more independence when they maintain skills across time (Schuler, 1995), most researchers did not address this area. Durand and Carr (1992) checked for maintenance with naïve trainers, but not with the children’s’ teachers or parents, which would also address generalization. Based on the evidence provided in this review, no research team conducted long term follow up studies to identify children who maintained low levels of challenging behaviors and high levels of communication. Further, research teams did not plan for maintenance across time. For example, Wacker et al. (1990) analyzed across topographies of behavior, but
did not plan for skill maintenance. Without following the participants and periodically checking for the use of taught skills, researchers may not know if the interventions aid children with autism in developing independence.

**Generalization**

Similar to the benefits of skill maintenance, children achieve greater independence when they generalize skills across settings and people (Layton & Watson, 1995). Generalization is particularly difficult for children with autism because they often remember tasks specific to the situation (Siegel, 1996). For example, Grandin (1995) described her experience as a young child and her insistence on routine. When a therapist taught her a task, she assumed the task applied to sessions with her therapist, thus, Grandin continued to engage in challenging behaviors in other settings. Additionally, children with autism frequently develop communication that only one other person recognizes (Schuler, 1995). A mother of a child with autism for example described a scenario where her child depended on her for a glass of water because she interpreted his grunting as a request (Maurice, 1993). When the child attended school, he screamed and hit himself when other individuals did not know he was thirsty. If researchers planned for generalization, children with autism likely would not exhibit such outbursts. Most researchers however continue conducting research without considering generalization. For example, only two of the eight research teams extended the research beyond the typical research environment. Further, only one of these studies occurred in a natural environment, that is, in a setting with people the child typically encounters.

**Implications for Researchers and Practitioners**

Since one of the goals of education is to improve the quality of life for the child and parents, FCT should occur in natural environments. Training in the natural environment teaches the child to associate the components of FCT with teachers, parents, classrooms, and home (i.e., generalization). Further, natural environments pose sensory issues (e.g., background noises, various lighting, other visual stimuli) for children with autism to overcome that is not present in stagnant environments. Therefore, future research should be directed at training teachers in the classroom and parents in the home.

**Training Teachers**

Training classroom teachers allows the child to associate FCT with the teacher and classroom. For example, if research teams teach a child in his or her classroom to ask for help completing a puzzle, the child will know to mand the request when performing the task at later times in the same room. For the process to be more beneficial, the teams should have the child’s teacher train him or her to mand the request. Therefore, the child will communicate with the teacher rather than rely on researchers who leave after the completion of the study.

After the research teams leave, teachers should continue to teach communication skills. Two ways teachers may enhance communicative behavior include: (a) taking advantage of naturally occurring opportunities and (b) arranging the environment to be conducive to communication.

**Natural opportunities.** During the school day, opportunities to teach communication skills abound. One period of particular interest is lunchtime because of the numerous communicative interactions naturally within the lunch routine. For example, as students progress through the lunch line, they choose a drink. Teachers may use this chance to teach students with autism to mand a request for milk or water. Similarly, teachers may use routines in the classroom to teach mands. For example, during coloring activities, the students may mand for markers or other desired tangibles.

**Arrange environment.** In addition to teaching mands during natural routines, teachers may arrange the classroom environment to encourage communication. For example, a teacher placing desired objects on shelves in view of, but out of reach of the children with autism creates a situation where a child desires to mand a request for an object. Further, teachers may include interests of children with autism in classroom activities and subse-
Training Parents

Besides training teachers, research teams also should focus on training parents to use FCT. Training parents in their home accomplishes two goals. First, the child will associate communication with the home environment. If a child needs help obtaining an item on a shelf in his or her bedroom, the parents train the child to request help in that setting. For greatest benefit, parents should train their child in each room of the home, allowing for generalization to all home settings. Second, the parents will begin to reinforce communication that others understand. For example, rather than giving the child water when he or she grunts, the parents who use FCT give the water to the child after he or she responds with the trained communication mand.

Parents may enhance communication in the home by using approaches similar to those teachers use in schools. That is, communication of children with autism may be enhanced in two ways: (a) by taking advantage of the natural environment, and (b) by arranging the environment to be conducive to communication.

Natural environment. The home includes several naturally occurring routines for increasing communication skills in children with autism. Mealtime is an excellent period to enhance communication. For example, children with autism may be taught to request certain food and more portions of the food. The key for the latter part is to give the children small portions, thus, increasing the opportunities for requesting.

Arrange environment. Parents may also arrange the home environment to increase opportunities to request. For example, when parents place their child’s favorite items out of reach, the child must request for the item. Further, parents may play on the floor with their child’s preferred toys, thus, creating another opportunity for their child to communicate.

Conclusion

The high prevalence and incidence of autism combined with problems in communication and behavior demands a response from the field of special education. Spanning the past twenty years, research teams responded with FCT. Most research on FCT however does not include children with autism. Further, the majority of research remains the same as Carr and Durand’s (1985) first published article regarding FCT, that is, clinically based. Future research teams should address maintenance and generalization by training teachers in classrooms and parents in homes while collecting data across time.

References


reports from my life with autism. New York: Vintage Books


Received: 2 June 2005
Initial Acceptance: 1 August 2005
Final Acceptance: 15 November 2005

224 / Education and Training in Developmental Disabilities-September 2006
Effectiveness of Simultaneous Prompting in Small Group: The Opportunity of Acquiring Non-target Skills through Observational Learning and Instructive Feedback

Oguz Gursel, Elif Tekin-Iftar, and Funda Bozkurt
Anadolu University

Abstract: A multiple probe study across behaviors, replicated across students, assessed the effectiveness of simultaneous prompting (SP) in a small group teaching arrangement on teaching (a) to show the provinces, rivers, and border countries of Turkey on a map and (b) to expressively identify the names of the symbols which are usually used in math. Subjects of the study were five middle school age students with developmental disabilities. Maintenance and generalization effects of SP were investigated in the study as well. Moreover, acquisition of non-target skills was also assessed through instructive feedback and observational learning. Results show that SP was effective. Students generalized and maintained the acquired behaviors. Assessment of observational learning and instructive feedback data showed that students acquired non-target skills to certain extents. Implications and future research needs are discussed.

One of the most desired strategies for teaching skills to students with disabilities is group teaching arrangement. Group teaching arrangement has some advantages over traditional one to one teaching arrangement on both teacher and student sides such as (a) less personnel and instruction time is needed, (b) students are placed in their least restrictive environment, (c) students have a chance to interact with their peers appropriately, (d) teachers provide instruction to more than one student at a time, (e) students have the advantage of observing their peers in the group and the chance of learning more (Collins, Gast, Ault, & Wolery, 1991). Group teaching arrangement has been widely used with considerable success on teaching both discrete and chained skills to students with disabilities (Alig-Cybriwsky, Wolery, & Gast, 1990; Doyle, Gast, Wolery, Ault, & Farmer, 1990; Parker & Schuster, 2002; Schoen & Sivil, 1989; Wolery, Cybriwsky, Gast, & Boyle-Gast, 1991).

One instructional procedure used to teach students with disabilities is known as simultaneous prompting (SP). In this procedure the teacher delivers the target stimuli and control ling prompt simultaneously. Therefore, the student does not have an opportunity to respond independently during training sessions and probe sessions are needed to test the transfer of stimulus control (Dogan & Tekin-Iftar, 2002; Gibson & Schuster, 1992; MacFarland-Smith, Schuster, & Stevens, 1993; Parrott, Schuster, Collins, & Gassaway, 2000; Tekin-Iftar, 2003; Tekin-Iftar, Acar, & Kurt, 2003).

To date there are 20 published studies examining the effects of SP on teaching either discrete or chained skills to people with various disabilities. Research has shown that SP is effective in teaching students with various disabilities such as moderate and severe mental retardation (Dogan & Tekin-Iftar, 2002; Fetko, Schuster, Harley, & Collins, 1999; Fickel, Schuster, & Collins, 1998; Maciag, Schuster, Collins, & Cooper, 2000; Parrott et al., 2000; Schuster & Griffen, 1993; Singleton, Schuster, & Ault, 1995); mild mental retardation (Palmer, Collins, & Schuster, 1999);
learning disabilities (Johnson, Schuster, & Bell, 1996), and developmental delays (Gibson & Schuster, 1992; MacFarland-Smith et al., 1993; Sewell, Collins, Hemmeter, & Schuster, 1998; Wolery, Holcombe, Werts, & Gipolone, 1993).

Evidence-based studies examined effectiveness of SP on teaching discrete tasks such as identifying occupations from picture cards (Dogan & Tekin-Iftar, 2002); object naming (MacFarland-Smith et al., 1993); science vocabulary words (Johnson et al., 1996); word identification (Griffen, Schuster, & Morse, 1998); community signs (Singleton et al., 1995; Tekin-Iftar, 2003); rebus symbols (Wolery et al., 1993); sight words (Schuster, Griffen, & Wolery, 1992; Gibson & Schuster, 1992); identifying national flags, stating the sums of addition facts, identifying unlabelled outlines of the states from the USA map, and demonstrating manual signs for communication picture symbols (Fickel et al., 1998); identifying animals (Tekin & Kircaali-Iftar, 2002); reading grocery aisle headers and occupational words, defining prefixes, identifying elements from Periodic Table (Parker & Schuster, 2002); verbal identification of manual signs (Palmer et al., 1999); identifying first-aid materials (Tekin-Iftar et al., 2003). SP was also used for teaching chained tasks such as making juice from frozen concentrate (Schuster & Griffen, 1993); dressing skills (Sewell et al., 1998); vocational skills (Fetko et al., 1999); construction of shipping boxes (Maciag et al., 2000), and hand washing (Parrott et al., 2000).

Small group instructional arrangement was used in only 35% of these 20 published studies. Homogeneous group format was used in the majority of studies. When delivering instruction with SP, using heterogeneous group is rare. According to Collins et al. (1991) conducting heterogeneous group may be more difficult as teacher will need to teach various skills in the group. However, students in this group have the chance of observing their peers in the group and may learn additional skills. Therefore, it can be said that conducting heterogeneous group may be more efficient than conducting homogeneous group or one to one teaching arrangement. Research has shown that students of various ages and ability levels can learn additional skills during group instruction through observational learning (Farmer, Gast, Wolery, & Winterling, 1991; McCurdy, Cundari, & Lentz, 1990; Parker & Schuster, 2002).

Instructive feedback is another instructional parameter that increases the number of behaviors learned during instructional trials. Werts, Wolery, Holcombe, and Gast (1995) defined instructive feedback as presenting extra, non-target stimuli, during consequent events of instructional trials. Students are not expected or reinforced to respond to these extra stimuli. Instructive feedback enhances efficiency of instruction by providing extra information during direct instruction. Werts et al. examined over 20 studies regarding presenting instructive feedback, and researchers reported that subjects gained some instructive feedback presented to them during instructional trials.

To date, there are only seven studies examining acquisition of instructive feedback while using SP delivered by either adults or peers (Griffen et al., 1998; Parrott et al., 2000; Schuster & Griffen, 1993; Singleton et al., 1995; Tekin-Iftar, 2003; Tekin-Iftar et al., 2003; Wolery et al., 1993). Among the seven studies examining the acquisition of instructive feedback during SP, four of them were conducted with elementary school students (Griffen et al.; Parrott et al.; Schuster & Griffen; Singleton et al.), one of them was conducted with preschool students (Wolery et al.), and one of them was conducted with middle school students (Tekin-Iftar et al.).

Findings of all above studies with SP showed that teachers implemented SP with high accuracy and most of these studies reported that SP is a relatively easy instructional procedure. However, there are only two studies investigating the effects of SP with middle school age students (Fickel et al., 1998; Tekin-Iftar et al., 2003). From these two, one study was conducted in a small group teaching arrangement and assessed the acquisition of observational learning (Fickel et al.). On the other hand there is no study investigating the effects of SP in small group on middle school age students and the acquisition of instructive feedback stimuli and observational learning stimuli in the group.

Therefore, the present study was conducted to examine effects of SP on teaching to show
the provinces, rivers, and border countries of Turkey on a map and to expressively identify the names of symbols that are frequently used in math. The following research questions were addressed in this study: (a) Is SP delivered in a heterogeneous small group effective on teaching to show the provinces, rivers, and border countries of Turkey on a map and expressively identify symbols which are frequently used in math for five students with developmental disabilities?, (b) Will students maintain the acquired behaviors over time (2 and 6 weeks after training)?, (c) Will students generalize the acquired behaviors across different persons and materials?, (d) Will students acquire instructive feedback stimuli provided to them on subsequent events after the correct responses during instructional trials?, (e) Will students acquire the target behaviors of their pairs through observational learning?

Method

Participants

Participants were selected by conducting interviews with their classroom teacher and parents at a public special school for students with developmental disabilities. The purpose of the study was shared with them. After obtaining their permissions, five students with developmental disabilities, three girls–two boys, were included in the study. All attended the same class at the same special school. None of them had a history with SP.

Prerequisite skills which students had to have were as follows: (a) attending to audio and visual stimuli for at least 10 minutes, (b) having turn taking skill, (c) following verbal instruction, (d) selecting reinforcers. All students had the prerequisite skills for this study. There was no adaptive behavioral score for the participants.

Giray (11 years 7 months old) functioned at mild to moderate range of intellectual disabilities. Areas of strength included self-care skills, fine and gross motor skills, receptive and expressive language skills. He had basic functional academic skills such as reading and writing. He had color, shape, and location concepts. Areas of weakness included reading comprehension and social skills.

Hale (13 years 1 month old) had Down syndrome. She functioned at moderate range of intellectual disabilities. She was receiving special education services since she was two and a half years old. Areas of strength included self-care skills, and fine and gross motor skills. She had knowledge of basic concepts such as color and shape concepts. Areas of weakness included functional academic skills and communication skills.

Sibel (14 years 3 months old) functioned at mild range of intellectual disabilities. Areas of strength included self-care skills, fine and gross motor skills, and receptive language skills. She could read and write, do addition and subtraction problems, and count exact change. She had knowledge of basic facts. Areas of weakness included communication, especially expressive language, and social skills.

Tarkan (12 years 2 months old) functioned at mild range of intellectual disabilities. Areas of strength included self-care skills, fine and gross motor skills, receptive and expressive language skills. He had basic functional academic skills such as reading and writing. He could do addition and subtraction problems with two digit numbers. Areas of weakness included social skills.

Irem (12 years 4 months old) functioned at mild range of intellectual disabilities as well. Areas of strength included self-care skills, fine and gross motor skills, receptive and expressive language skills. She had basic functional academic skills such as reading and writing. She could use phone and public transportation, and read a clock independently. She had color, shape, and location concepts. She could do addition and subtraction problems with two digit numbers. Areas of weakness included social skills.

Dyads were formed to assess acquisition of observational learning stimuli. Giray and Hale, Sibel and Tarkan, and Irem and Sibel were the first, second and third dyads respectively. Since five students participated in the study Sibel was paired with two different students. In the last dyad, acquisition of the observational learning was assessed only for Irem since Sibel’s acquisition was tested in the second dyad. The third author conducted all experimental sessions. She had a master’s degree in special education and four years experience in teaching students with intellectual
disabilities. Reliability data were collected by the first author who is faculty at the Department of Special Education at Anadolu University.

Setting

The study was conducted in the students’ classroom (5 m x 3 m). There was a rectangular teacher table, chairs for the students, several tables for the students, and a board in the classroom. Students and researcher sat down face to face at a table in a semi circle. All experimental sessions were conducted in the same classroom. Intervention sessions were conducted in group teaching arrangement, and the rest of the experimental sessions were conducted in one to one teaching arrangement. Intervention and probe sessions were conducted Monday, Wednesday and Thursday at 10:00 to 10:30 am. The researcher recorded each session via camcorder. No one was available during the experimental sessions other than the researcher.

Materials

During training, index cards (5 cm x 5 cm), maps, reinforcers, a camcorder, and a stopwatch were used. Index cards were used to teach symbols used in math (e.g., min, gr, /). Reinforcers were selected by the students and consisted of objects such as stationery items and music tapes. Nine index cards were used when teaching the symbols. Each card had a single symbol. Sixteen point Times New Roman font was used in the cards. Nine symbols were chosen to teach one student. Three training sets of symbols were formed for the student. Reference map (45 x 60 cm) was used when teaching to show the provinces and border countries of Turkey, and environmental map (45 x 60 cm) was used when teaching to show the rivers in Turkey. Furthermore, as generalization items, different maps on different sizes and index cards in different sizes and colors were used.

Selection of Target Behaviors

Target behaviors were selected from IEP’s of each student. They were selected from two curriculum areas: Social Sciences and Math. The rivers, provinces and border countries of Turkey were taken from “Our Country and Our Regions” unit of Social Sciences class and symbols were taken from various units of Math class. Target behaviors were defined as “when asked student shows the border countries (rivers or provinces) of Turkey on a map” and “when shown student tells the name of the symbol on a card.”

Screening Sessions

Screening sessions were conducted individually to identify the prospective target stimuli for each student. Prior to initial baseline conditions, 35 provinces, five provinces from seven regions in Turkey, were selected to form a pool. A pool for the rivers, 18 rivers, and a pool for the math symbols, 15 symbols, were formed. After that, to identify the unknown stimuli from these polls for each student, two consecutive screening sessions were conducted with a trial for each prospective target stimuli. The trials were presented in a random order. Instructive feedback stimuli were also screened in the same sessions.

Screening sessions were conducted as follows. The teacher had the materials ready, and secured the students attention (e.g., “Sibel, lets start to work with you. Are you ready?”). After receiving an affirmative response, the teacher presented the task direction, (e.g., “Sibel, please show Bursa on the map?”), and waited 4 s. After waiting 4 s, the teacher asked the instructive feedback stimuli, the region where that province is from, (e.g., “Sibel, tell me which region is Bursa from?”). Correct and incorrect responses for target behaviors as well as responses for instructive feedback were ignored during the screening sessions. The nine stimuli (provinces, rivers, border countries, and symbols) to which the students did not respond correctly were chosen as target behaviors. Three training sets were prepared for each student and each training set had three target behaviors. Target behaviors were randomly assigned to training sets. The target behaviors, and training sets for the dyads and the instructive feedback presented with each target behavior are in Table 1 and Table 2 respectively.

Screening sessions of observational learning stimuli were conducted after forming the
training sets of each student. The students were tested in the same manner about the target behaviors of their pair in the dyad.

The students’ attention and cooperation were reinforced verbally at the end of each session (e.g., “Very good Sibel. You paid attention and were cooperative with me today.”).

### General Procedures

Screening sessions were conducted to identify target behaviors prior to the experimental procedures. Nine target behaviors were taught to each student in three training sets. All sessions were conducted and recorded by the third author. During instructional trials, instructive feedback was delivered after each correct response. Observational learning was encouraged and reinforced during instructional trials as well. Full and daily probe sessions were conducted. Also, maintenance probe sessions for targeted behaviors, instructive feedback and observational learning probe sessions, and generalization probe sessions across persons and materials were conducted. Instructional sessions were conducted in small group teaching arrangement and the rests of the other experimental sessions were conducted in one to one teaching arrangement. Individual criteria were used during training. Response intervals and intertrial intervals during all experimental sessions were 4 s. Students received verbal reinforcement for their attending and cooperation behaviors at the end of all sessions by the teacher.

### Full Probe Conditions

Full probe sessions were conducted in one to one teaching arrangement before introducing the intervention to the first training set and after criterion were met for each training set. All training sets were probed during full probe sessions until stable data were recorded for at least three consecutive sessions. Each stimulus in the training sets for each student was presented three times during the sessions. The teacher randomly sequenced the stimuli before the sessions. Full probe sessions were implemented as follows: the teacher had training materials ready, secured the student’s attention (e.g., “Are you ready?”), and then provided the target stimulus and waited 4 s for the student to respond. The teacher recorded the student’s responses and correct responses resulted in verbal praise; incorrect or no responses were ignored.

### Daily Probe Conditions

Since a controlling prompt was delivered on every training trial, the student did not have an opportunity to respond to the target stimulus independently. Therefore, daily probe sessions were conducted to test for transfer of stimulus control in SP. Daily probe sessions were conducted before every single daily training session. Training sets that were currently being taught were probed in these sessions. No daily probe session was conducted before the first training session. Correct responses during daily probe sessions were counted toward criterion. Criterion was 100% correct responding for three consecutive daily probe sessions. Daily probe sessions were implemented just like full probe sessions with one exception. Only the currently trained set was assessed in the daily probe session. Same sequences were provided in daily probe sessions.

### Instructive Feedback and Observational Learning Probe Sessions

Following every single full probe condition, instructive feedback and observational learn-
TABLE 2
Training Sets and Instructive Feedback Stimuli

<table>
<thead>
<tr>
<th>Dyads Participants First</th>
<th>Training Sets</th>
<th>Instructive Feedback Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyad (Giray-Hale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giray</td>
<td>1 Georgia</td>
<td>Tbilisi is the capital of Georgia.</td>
</tr>
<tr>
<td></td>
<td>2 Armenia</td>
<td>Yerevan is the capital of Armenia.</td>
</tr>
<tr>
<td></td>
<td>3 Azerbaijan</td>
<td>Baku is the capital of Azerbaijan.</td>
</tr>
<tr>
<td></td>
<td>2 Iran</td>
<td>Teheran is the capital of Iran.</td>
</tr>
<tr>
<td></td>
<td>3 Greece</td>
<td>Athens is the capital of Greece.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hale</td>
<td>1 Bursa</td>
<td>Bursa is in the Marmara region in Turkey.</td>
</tr>
<tr>
<td></td>
<td>2 Batman</td>
<td>Batman is in the Eastern Anatolian region in Turkey.</td>
</tr>
<tr>
<td></td>
<td>3 Corum</td>
<td>Corum is in the Black Sea region in Turkey.</td>
</tr>
<tr>
<td></td>
<td>2 Aydin</td>
<td>Aydin is in the Aegen region in Turkey.</td>
</tr>
<tr>
<td></td>
<td>3 Samsun</td>
<td>Samsun is in the Black Sea region in Turkey.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Dyad (Sibel-Tarkan)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibel</td>
<td>1 Izmir</td>
<td>Izmir is in the Aegen region in Turkey.</td>
</tr>
<tr>
<td></td>
<td>2 Maras</td>
<td>Maras is in the Mediterranean region in Turkey.</td>
</tr>
<tr>
<td></td>
<td>3 Istanbul</td>
<td>Istanbul is in the Marmara region in Turkey.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tarkan</td>
<td>1 S</td>
<td>It is used when telling time.</td>
</tr>
<tr>
<td></td>
<td>&gt;</td>
<td>It is used when talking about greater than.</td>
</tr>
<tr>
<td></td>
<td>Kg</td>
<td>It is used when measuring weight.</td>
</tr>
<tr>
<td></td>
<td>2 Cm</td>
<td>It is used for when measuring height and length.</td>
</tr>
<tr>
<td></td>
<td>±</td>
<td>It is used in the subtraction problems.</td>
</tr>
<tr>
<td></td>
<td>3 Min</td>
<td>It is used when telling time.</td>
</tr>
<tr>
<td></td>
<td>Gr</td>
<td>It is used when measuring weight.</td>
</tr>
<tr>
<td></td>
<td>/</td>
<td>It is used in division problems.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third dyad (Irem-Sibel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irem</td>
<td>1 Sakarya</td>
<td>Sakarya flows into Black Sea.</td>
</tr>
<tr>
<td></td>
<td>2 Kizilirmak</td>
<td>Kizilirmak flows into Black Sea.</td>
</tr>
<tr>
<td></td>
<td>3 Yesilirmak</td>
<td>Yesilirmak flows into Black Sea.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ing probe sessions occurred in order to assess the acquisition of the stimuli introduced in the instructive feedback and observational learning trials. Nine trials occurred for each student during these sessions. These sessions were conducted in the same format with full probe sessions.

The teacher had the materials ready, secured the student’s attention, and presented the task direction, “... tell me, which region is Bursa from?” and waited for 4 s. There were correct responses, incorrect responses, and no responses during instructive feedback probe sessions. Correct responses were defined as telling the region of the asked province correctly within 4 s. Incorrect responses and no responses were defined as telling a different region or not responding within 4 s. Correct responses resulted in verbal descriptive praise, incorrect responses and no responses were ignored, and the next trial was presented. Students received verbal reinforcement for their attending and cooperation behaviors during the sessions.

Acquisition of the observational learning stimuli was assessed in the dyad by testing the target stimuli of the student’s pair on the other student in the study (e.g., Hale was tested about Giray’s target stimuli and Giray was tested about Hale’s target stimuli). The teacher conducted observational learning probe sessions as follows: The teacher had materials ready, secured the student’s attention and presented the task direction, “Hale, please show Syria on the map.” The possible responses of the students and their consequences were the same as instructive feedback probe sessions.

**Simultaneous Prompting Procedure**

After obtaining consistent data during baseline sessions, the teacher started to use SP to teach target behaviors to students in small group arrangement. Simultaneous prompting and instructive feedback were delivered during instruction to show the provinces, rivers, and border countries of Turkey on the map, and to expressively identify the symbols that are usually used in math. Training was delivered three days a week with one training session each day. There were nine trials for each student. Each target behavior in the training sets was presented three times randomly. A total of 45 trials were delivered with nine trials for each student in the group. Prior to each training session, the teacher determined the order of presenting the trials and of starting with the student. Responses during instruction with SP were scored as correct, incorrect, and no response. Responses were defined the same as in the probe sessions. Different controlling prompts were used in the study for the students.

Training sessions were conducted as follows. The teacher had the materials ready, and secured the students’ attention in the group by delivering attention cue for the group. After receiving an affirmative response to the question, “Students, are you ready for work?”, the teacher explained the group that she was going to work with one of them and everybody in the group needed to listen carefully especially his/her pair in the group (i.e., “I’m going to start with Giray today. All of you should observe us carefully, especially his/her Hale.”). After that the teacher presented the task direction, “Giray, Please show Bulgaria on the map,” and then provided the controlling prompt immediately, “teacher showed Bulgaria on the map”, and waited 4 s for a response. If the students imitated the controlling prompt and repeated it within 4 s, the teacher provided a verbal reinforcement “Very good, Giray. You show Bulgaria on the map.” and then provided the instructive feedback “The capital of Bulgaria is Sofia.” Incorrect responses or no responses within 4 s resulted in reproducing the controlling prompt and the teacher presented the next trial. Students’ attention and their cooperation behaviors were reinforced at the end of the sessions by the teacher.

Since acquisition of the observational learning stimuli was one of the parameters in the study, observational learning was encouraged during the training. The teacher secured the other students’ attention while working with one of the students in the group. The teacher verbally reinforced their observing behaviors during the session (e.g., “Good job. You all observe Giray very good.”). Continuous reinforcement schedule was used until criterion was met, then reinforcements were delivered on a VR5 basis.
Generalization across persons and materials probe sessions was conducted in one to one teaching arrangement in a pretest-posttest manner. These sessions occurred before any training as a pretest, and at the end of teaching all training sets, final full probe session, as a posttest. Maintenance probe sessions were conducted two and six weeks after training, following the final full probe session. Generalization and maintenance probe sessions were conducted just like full probe sessions. However, generalization sessions were conducted by the first author and different maps and index cards were used during the sessions. Reinforcement was thinned (i.e., VR3 for the first probe session, FR9 for the consecutive session) during maintenance and generalization sessions.

Experimental Design

A multiple probe design across training sets and replicated across students was used to investigate effectiveness of SP delivered in small group teaching arrangement on teaching to show the provinces, rivers, and border countries of Turkey and to expressively identify the name of a given symbol used in math frequently to students with developmental disabilities. The dependent measure was percentage of correct responses on showing the provinces, rivers, and border countries of Turkey and expressively identifying the name of a given symbol which are used in math frequently, and the independent variable of the study was SP. The independent variable was introduced to one training set at a time. Experimental control was built in when the student was responding at or near to baseline levels during full probe conditions before the intervention had been introduced and the criterion was reached only after the intervention was introduced (Tekin-Iftar & Kircaali-Iftar, 2004; Wolery, Bailey, & Sugai, 1988).

Interobserver and Procedural Reliability

Reliability data were collected at least 20% of all experimental sessions (20% of full probe and daily probe sessions, 20% of training sessions; 33% of maintenance sessions and instructive feedback; and 50% of generalization sessions). A point by point method with a formula of the number of agreements divided by the number of agreements plus disagreements multiplied by 100 was used to calculate interobserver reliability (Tawney & Gest, 1984; Tekin-Iftar & Kircaali-Iftar, 2004). Interobserver reliability data collected during the full probe sessions yielded a percentage of agreement of 100% across all students. Dependent measure reliability data collected during daily probe sessions yielded a mean percentage of agreement of 99.3% (range = 89 – 100), and 100% during training sessions across all students. Dependent variable reliability data collected during instructive feedback, observational learning, maintenance, and generalization probe sessions yielded a mean percentage of agreement of 98.9% (range = 92 – 100), 99.5% (range = 95 – 100), 95% (range = 84 – 100), 98.2% (range = 78 – 100) respectively across all students.

Procedural reliability data were collected to estimate whether the teacher delivered SP and other experimental sessions (e.g., full and daily probe sessions, generalization and maintenance sessions, instructive feedback and observational learning probe sessions etc.) as they were planned in the study. Planned steps that the teacher was expected to demonstrate during simultaneous prompting sessions were (a) having materials ready, (b) securing the student’s attention, (c) encouraging observational learning, (d) presenting task direction, (e) providing controlling prompt immediately after the task direction, (f) delivering correct consequences, (g) delivering correct instructive feedback, and (h) providing appropriate inter-trial interval (4 s). Planned steps that the teacher was expected to demonstrate for daily, full, instructive feedback and observational learning, and generalization and maintenance probe sessions were (a) having materials ready, (b) securing the student’s attention, (c) presenting the task direction, (d) delivering correct consequences, and (e) providing the appropriate inter-trial interval (4 s). Procedural reliability was calculated by dividing the number of observed teacher behaviors by the number of planned teacher behaviors, and multiplied by 100 (Billingsley, White, & Munson, 1980; Tekin-Iftar & Kircaali-Iftar, 2004; Wolery, Bailey, & Sugai, 1988).
Independent variable reliability data indicated that the teacher performed all behaviors with 100% accuracy during all probe sessions. During training sessions, the teacher implemented all behaviors with 100% accuracy with the exception of delivering instructive feedback stimuli. She delivered instructive feedback stimuli with a mean of 83% accuracy (range = 67 – 100) across all students.

Results

Instructional Data

Figures 1–5 display the percentage of correct responses during full probe, daily probe and maintenance probe sessions for Giray, Hale, Sibel, Tarkan, and Irem respectively. As seen in the figures, using SP to teach a heterogeneous group of students with developmental disabilities was effective. Any procedural modification was not needed during the experimental sessions. Hale did not attend school during training with her third training set. Number of training sessions and trials, training and probe time, and training and probe errors are presented in Table 3.

Sessions and Trials Through Criterion

Seventy-eight training sessions and 702 training trials were needed for the students to meet criterion on all training sets. Giray needed 21 training sessions and 189 training trials, Hale needed 14 training sessions and 126 training trials, Sibel needed 12 training sessions and 108 training trials, Tarkan needed 15 training sessions and 135 training trials, and Irem needed 16 training sessions and 144 training trials. Giray needed the highest number of training sessions through criterion and Sibel needed the lowest in the group.

Training and Probe Time Through Criterion

Two hr, 32 min, 23 s training time was needed through criterion across students. Giray, Sibel, Tarkan, Irem needed 45 min, 8 s, 22 min 13 s, 25 min 42 sec, and 35 min 2 s training time through criterion across all training sets respectively. Hale needed 24 min 18 s training time though criterion across first two training sets. The training time that the students needed through criterion was between 22 min 13 s and 45 min 8 s. 1 hr 11 min, 6 s probe session time was needed across five students through criterion. The individual probe time across training sets were between 9 min 25 s and 17 min 39 s.

Training and Probe Error Through Criterion

SP instructional sessions were almost errorless for the students. One error occurred during training sessions with Sibel, 2 with Irem, 3 with Hale, 4 with Giray, and 5 with Tarkan. Fifteen errors occurred during training with 2.14%. There were 172 errors during probe sessions with an average of 6.57% across students. Probe session error rate ranged from 0% to 57.7%.

Maintenance and Generalization

Maintenance probe sessions were conducted two and six weeks after the final full probe sessions. Maintenance data for the students showed that students maintained the acquired skills of showing the provinces, rivers, and border countries of Turkey on a map and expressively identifying the names of the symbols which are frequently used in math at criterion level (see Figures 1-5). Generalization across persons and materials data showed that except Irem all students generalized the acquired skills at criterion level. Irem generalized the acquired skill at 56% across persons and materials. Pretest generalization measures across sets were 0% for Giray, Tarkan, Hale, and Irem whereas posttest generalization measures across all sets were 85% for Giray, and 100% for Sibel, Tarkan, and Hale.

Instructive Feedback Data

Data collected indicated that each student in the group acquired some of his/her own instructive feedback stimuli. Mean percentage of correct responding on instructive feedback stimuli for the training set for each student during screening, full probe and maintenance sessions are presented in Table 4. During baseline all students’ responses were at 0%.
correct responding. When experimental sessions were over (after the final probe session) the acquisition of the instructive feedback across training sets was between 33% and 100%.

Observational Learning Data
Data collected for the acquisition of observational learning indicated that students acquired some of the target behaviors of their
pairs to a certain extent by observational learning. Mean percentage of correct responding on observational learning stimuli for the training set for each student during screening, full probe and maintenance sessions are presented in Table 5. During baseline students’ responses were between 0%–33% correct responding. When experimental
sessions were over (after the final probe session) the acquisition of the observational learning stimuli across training sets was between 33% and 100%.

Discussion

The purpose of this study was to evaluate the effects of SP delivered in a small group on
teaching to show the provinces, rivers and border countries of Turkey on a map and to expressively identify the names of the symbols, which are frequently used in math to five students with developmental disabilities. Generalization and maintenance effects of SP were examined as well. In addition, acquisition of instructive feedback stimuli and observational learning stimuli were investigated in the study. Based on the data col-

Figure 4. Percent of correct responses during full, daily and maintenance probe sessions for Tarkan.
lected, several findings and implications are worth to discuss.

First, the data indicated that SP delivered in small group was effective on teaching to show the provinces, rivers, and border countries of Turkey on a map and to expressively identify the name of the symbols which are frequently used in math to five students with developmental disabilities. Findings of the study are consistent with the findings of the previous studies. As mentioned before most published studies with SP were designed to teach dis-

Figure 5. Percent of correct responses during full, daily and maintenance probe sessions for Irem.
crete behaviors such as science vocabulary words (Johnson et al., 1996), object naming (MacFarland-Smith et al., 1993; Tekin-Iftar et al., 2003), word identification (Griffen et al., 1998; Schuster et al., 1992), community signs (Singleton et al., 1995; Tekin-Iftar, 2003; Wolley et al., 1993); and animal identification (Tekin & Kircaali-Iftar, 2002), identifying national flags, stating the sums of addition facts, identifying unlabelled outlines of the states from the US map, and demonstrating manual signs for communication picture symbols (Fickel et al., 1998). Very few of them were conducted in group teaching format (Fickel et al; Palmer et al., 1999; Parker & Schuster, 2002). The findings of the present study are consistent with the findings of these studies. Therefore, it can be claimed that the present study extends current literature about the effectiveness of SP when delivered in small group.

Second, data indicated that students were able to maintain the acquired behaviors over time (i.e., 2 and 6 weeks after training). These findings are also consistent with the findings of the previous studies. However, maintenance data were collected for only three students. Student attrition (i.e. Hale) and starting of the summer holiday (i.e., Tarkan) were the main reasons for this limited findings. Third, it was observed that students generalized the acquired behaviors across persons and materials to a certain extent. The generalization range for students was between 56% and 100% for the students. Therefore, it can be argued that generalization effects of SP were positive in general. These findings are also consistent with the previous studies.

Fourth, data showed that students in the group gained some of the instructive feedback stimuli presented to them on the consequent events during instructional trials. As mentioned earlier an efficient instructional procedure allows students learn extra stimuli during training. In other words an efficient instructional procedure increases the number of behaviors learned during instructional trials. From this perspective, efficacy of SP can be

<table>
<thead>
<tr>
<th>Student/Set</th>
<th>No. training sessions</th>
<th>No. training trials</th>
<th>No. training errors</th>
<th>% training errors</th>
<th>Training time</th>
<th>Daily probe time</th>
<th>No. probe errors</th>
<th>% probe errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gökhan</td>
<td>1</td>
<td>12</td>
<td>108</td>
<td>3</td>
<td>2.7</td>
<td>24 min 16 s</td>
<td>10 min</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>8 min 4 s</td>
<td>2 min 39 s</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6</td>
<td>54</td>
<td>1</td>
<td>1.85</td>
<td>12 min 48 s</td>
<td>5 min</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>189</td>
<td>4</td>
<td>1.5</td>
<td>45 min 8 s</td>
<td>17 min 39 s</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Hale</td>
<td>1</td>
<td>11</td>
<td>99</td>
<td>2</td>
<td>2.02</td>
<td>18 min 13 s</td>
<td>11 min 55 sn</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>27</td>
<td>1</td>
<td>3.70</td>
<td>6 min 5 s</td>
<td>3 min</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>126</td>
<td>3</td>
<td>2.86</td>
<td>24 min 18 s</td>
<td>14 min 55 s</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Sibel</td>
<td>1</td>
<td>5</td>
<td>45</td>
<td>1</td>
<td>2.2</td>
<td>8 min 32 s</td>
<td>3 min</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>6 min 21 s</td>
<td>3 min</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>7 min 30 s</td>
<td>2 min 30 s</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>108</td>
<td>1</td>
<td>.73</td>
<td>22 min 23 s</td>
<td>9 min 25 s</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Tarkan</td>
<td>1</td>
<td>7</td>
<td>63</td>
<td>4</td>
<td>6.34</td>
<td>10 min 30 s</td>
<td>5 min</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
<td>45</td>
<td>1</td>
<td>2.2</td>
<td>9 min 27 s</td>
<td>6 min 10 s</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>5 min 45 s</td>
<td>2 min 24 s</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>135</td>
<td>5</td>
<td>2.85</td>
<td>24 min 42 s</td>
<td>13 min 49 s</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Irem</td>
<td>1</td>
<td>9</td>
<td>81</td>
<td>2</td>
<td>2.46</td>
<td>18 min 39 s</td>
<td>8 min 24 s</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>9 min 16 s</td>
<td>4 min</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>7 min 7 s</td>
<td>2 min 54 s</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>144</td>
<td>2</td>
<td>.83</td>
<td>35 min 2 s</td>
<td>15 min 18 s</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>78</td>
<td>702</td>
<td>15</td>
<td>6.57</td>
<td>2 h 30 m 23 s</td>
<td>1 h 11 min 6 s</td>
<td>172</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 3
Instructional Data for Each Student and Training Set Through Criterion

Simultaneous Prompting in Small Group / 239
### TABLE 4
Accuracy of Responding to Instructive Feedback During Full Probe Conditions

<table>
<thead>
<tr>
<th>Tutees</th>
<th>Sets</th>
<th>Screening</th>
<th>Probe I</th>
<th>Probe II</th>
<th>Probe III</th>
<th>Probe IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giray</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>33%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>Hale</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>-</td>
</tr>
<tr>
<td>Sibel</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>Tarkan</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Irem</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>84%</td>
<td>89%</td>
</tr>
</tbody>
</table>

### TABLE 5
Accuracy of Responding to Observational Learning During Full Probe Conditions

<table>
<thead>
<tr>
<th>Tutees</th>
<th>Sets</th>
<th>Screening</th>
<th>Probe I</th>
<th>Probe II</th>
<th>Probe III</th>
<th>Probe IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giray</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>44%</td>
<td>56%</td>
<td>-%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>11%</td>
<td>44%</td>
<td>100%</td>
<td>-%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>11%</td>
<td>67%</td>
<td>100%</td>
<td>-%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>11%</td>
<td>32%</td>
<td>85%</td>
<td>-%</td>
</tr>
<tr>
<td>Hale</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>28%</td>
</tr>
<tr>
<td>Sibel</td>
<td>1</td>
<td>0%</td>
<td>33%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
<td>100%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>33%</td>
<td>67%</td>
<td>74%</td>
<td>100%</td>
</tr>
<tr>
<td>Tarkan</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>11%</td>
<td>0%</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>22%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>11%</td>
<td>0%</td>
<td>29%</td>
</tr>
<tr>
<td>Irem</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>78%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>67%</td>
</tr>
<tr>
<td>Total</td>
<td>Across Sets</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>78%</td>
<td>89%</td>
</tr>
</tbody>
</table>
seen clearly. To date, there are only seven studies examining the acquisition of instructive feedback during SP delivered by either adults or peers (Griffen et al., 1998; Parrott et al., 2000; Schuster & Griffen, 1993; Singleton et al., 1995; Tekin-Iftar, 2003; Tekin-Iftar et al., 2003; Wolery et al., 1993). The findings of the present study are consistent with these studies. Majority of the studies investigating the effects of SP were conducted with preschool and elementary school students, and one of them was conducted with middle school students (Tekin-Iftar et al.). In the present study students acquired the instructive feedback with 33% to 100% accuracy. This study contributes and enhances the current literature on delivering instructive feedback during SP trials to middle school age students with developmental disabilities.

Fifth, data collected for the acquisition of observational learning indicated that students acquired some of the target behaviors of their pairs to a certain extend by observational learning. The highest correct responding during baseline was 33% whereas, when experimental sessions were over, the acquisition of the observational learning stimuli across training sets was between 33% and 100%. These findings are consistent with the findings of the previous studies (Farmer et al., 1991; Fickel et al., 1998; McCurdy et al., 1990; Parker & Schuster, 2002).

Besides these findings several points observed during study are important to discuss. First, both observations of the records of training and procedural reliability data showed that although it was her first experience with SP in the group, the teacher implemented the SP in the group with high accuracy. This finding encourages us for advising professionals to use SP either in group or one to one teaching arrangement. Second, error rate during probe sessions was high like in previous studies. The error rate during daily probe sessions was consistently higher than the error rate during training sessions in the previous studies as well. On the other hand conducting daily probe sessions for five students was cumbersome for both students and the teacher. Therefore, several strategies can be advised to decrease the error rate during probe sessions and to deal with the effects of being continuously measured. Conducting intermittent probe sessions and delivering error correction during probe sessions can be taken into consideration as strategies for decreasing the error rate. Conducting intermittent probe sessions can also be helpful for dealing with the effects of being continuously measured. Future research should examine the effects of conducting different probe schedules and delivering error correction during probe sessions to deal with the above problems.

Although findings of the study were very encouraging the results should be interpreted cautiously for the following reasons. First, this study was limited with five students and teaching discrete skills. Use of SP with a larger group of students from various disability areas is warranted. Second, experimental control with Sibel could not be demonstrated in the study. Sibel was living in the orphanage and her sisters provided exercises to her about the target behaviors of the second training set of the study upon her request. Therefore, source of the progress of Sibel during second training set can not be solely explained by the effects of SP. The effects may be due to SP alone, or practice at home or both. Third, although different stimuli and different tasks were used in the study during training some of the students had the common stimuli and tasks (i.e., Hale and Sibel-showing provinces of Turkey on the map). Each student may have different stimuli and different tasks in the future studies. Fourth, the diagnoses of the students were mild and moderate intellectual disabilities. Also, the ages of the students were close to each other. Therefore, the results are limited with these features. While forming the groups, more heterogeneous groups, students who are diagnosed with different labels and vary in ages can be included in future research studies.

In addition to the above mentioned future research implications, the following research suggestions can be made when results of the study are taken into consideration. Future research should be conducted to examine similar effects when teaching chained skills with SP delivered in the small group. Massed trial presentation format was used in the study. The effects of other trial presentation formats such as, distributed and spaced, can be investigated in the future studies. Individual responding was utilized in the study. Future re-
searchers may design a study to investigate the effects of choral responding and/or compare both regarding the effectiveness and efficiency when delivering instruction with SP. Individual criterion was used in the study. However, group criterion is an alternative approach. The effects of using group criterion can be examined in the future studies. Also, comparison studies can be designed to investigate the differences between them, if any. Independent group contingency, each student received reinforcement based on his/her own behaviors, was used in the study. Interdependent and dependent contingencies can be taken as alternative parameters to investigate in the future research. Literature shows that peer tutor can deliver training with SP reliably in one to one teaching arrangement (Tekin-Iftar, 2003). Conducting training with SP in small group by the peers can be examined in the future research. Furthermore, future research might be designed to compare the effects of peer-delivered and teacher-delivered SP in terms of effectiveness, efficiency, and social validity variables in small groups.

References


Received: 12 April 2005
Initial Acceptance: 1 June 2005
Final Acceptance: 1 September 2005
Increasing Opportunities for Requesting in Children with Developmental Disabilities Residing in Group Homes through Pyramidal Training

Ralf W. Schlosser
Northeastern University

Elizabeth Walker
Bloorview MacMillan Centre

Jeff Sigafoos
University of Tasmania

Abstract: This study evaluated the effectiveness of pyramidal training with direct care staff in group homes. Training focused on teaching staff how to provide opportunities for communication to non-speaking children with developmental disabilities to communicate. Staff were taught through a combination of a workshop, and nondirective consultation following an adult education model. A multiple-probe design across three cohorts of direct care staff was used to evaluate the effectiveness of the training in terms of the number of opportunities provided by staff and the percentage of requests emitted by participating children. Results showed that pyramidal training resulted in more opportunities provided to the children and concurrent increases in children’s requesting. Results are discussed in terms of the suitability of pyramidal training for group home settings, and implications for future research.

Children with developmental disabilities and severe communication impairments frequently rely on augmentative and alternative communication (AAC) systems such as manual signs and/or aided systems such as communication boards and voice output communication aids to communicate (Beukelman & Mirenda, 1998; Lloyd, Fuller, & Arvidson, 1997; Reichle, York, & Sigafoos, 1991). Communication is considered a transactional process where partners influence each other in the course of the exchange (Light, Datillo, English, Gutierrez, & Hartz, 1992). Studies examining the interaction patterns between persons with developmental disabilities and communication partners in a variety of settings, however, indicate that partners exhibit high rates of directives while individuals with developmental disabilities are provided with relatively few opportunities for communication (Blindert, 1975; Houghton, Bronicki, & Guess, 1987; Sigafoos, Roberts, Kerr, Couzens, & Baglioni, 1994). In light of these interaction patterns, several authors have advocated that AAC intervention should be dual pronged, involving intervention with the individual using AAC and instruction of communication partners (Beukelman & Mirenda; Calculator & Luchko, 1983; Culp & Carlisle, 1988; Cumley & Beukelman, 1992;Naughton & Light, 1989; Walker & Antonius, 1995). A few controlled studies have begun to evaluate the efficacy of communication partner instruction in promoting increased communication by AAC users with developmental disabilities (Light et al.; Sack, McLean, McLean, & Spradlin, 1992; Sigafoos, Kerr, Roberts, & Couzens, 1994). For example, Light et al. successfully instructed three partners of two young adults with developmental disabilities who use AAC to decrease rates of turn-taking and initiations and to increase rates of turns that were responsive. As a result, the two adults increased the frequency of their initiations.

The need to provide instructions to partners is particularly evident in teaching an initial request repertoire to the learner (Sack et al., 1992; Sigafoos, Roberts et al., 1994). An initial requesting repertoire provides a direct...
benefit to the learner by allowing access to preferred objects and activities which require the mediation of another person (Sigafoos & Reichle, 1992; Skinner, 1957). This mediation of another person (i.e., the partner) requires the provision of opportunities so that the child can learn to request spontaneously instead of only when prompted. In a study by Sack et al., staff in a residential setting were taught to follow a scripted routine in order to provide specific requesting opportunities for five adolescent learners with severe mental retardation in a snack activity and an art activity. Results revealed that staff provided more opportunities and the learners increased their requests within the scripted routine context.

In another relevant study, Sack et al. (1994) taught five teachers serving 26 children with moderate to severe mental retardation to increase requesting opportunities. Due to this instruction, the opportunities provided by teachers and the requesting behavior of the children increased during intervention, and was maintained during follow-up.

A number of strategies are available for communication partners in order to create opportunities for requesting among persons with developmental disabilities (for a detailed review see Sigafoos, Roberts et al., 1994). The missing-item format, for example, involves withholding a needed object until the child requests it or attempts to request it independently or with prompting by the partner (Cipani, 1988). In the study by Sack et al. (1992) staff were taught to use the missing-item format. The interrupted-chain strategy represents another procedure (Goetz, Gee, & Sailor, 1985; Hunt, Goetz, Alwell, & Sailor, 1986). Here, an activity is interrupted to create an opportunity for the child to request, independently or with prompting, the continuation of the activity. Delayed assistance may be used as yet another strategy to create opportunities for requesting (Reichle, Anderson, & Schermer, 1986). When a child is noticed to struggle completing an activity, the partner may approach, but wait until the child requests “help” either independently or prompted by the partner. Sigafoos, Kerr and colleagues (1994) taught each of these three strategies to their participating partners. These strategies are often combined with delayed prompting; a technique that provides guidelines for the use of successively more and more intrusive partner prompts (Halle, Baer, & Spradlin, 1981; Halle, Marshall, & Spradlin, 1979). The partners in the Sack et al. and the Sigafoos, Kerr et al. study were also taught to use delayed prompting to solicit requests.

In addition to the content of partner instruction, instructors must also consider the effectiveness and suitability of available instructional formats to deliver this content. Sack et al. (1992) used an overview session, a video-taped demonstration of delayed prompting, followed by actual practice, supervised application in the activities targeted for training, and a review session. Sigafoos, Kerr et al. (1994) successfully employed a nondirective consultation format (Peck, Killen, & Baumgart, 1989) to instruct the teachers. This included an overview session, a session to generate ideas on how to use the strategies taught in the overview session, and feedback following each application. Nondirective consultation may represent an effective and acceptable form of teacher-directed intervention because it involves teachers in the training process and because it exploits the teacher’s greater familiarity with the children and existing routines (Sigafoos, Kerr et al). In absence of any research on the comparative effectiveness of these instructional formats, it is important to choose a training format that seems at least conceptually appropriate to the targeted partners. Because many communication partners in AAC are adults, efforts to prepare these adults should consider the characteristics of adult learners and models of adult learning (Cumley & Beukelman, 1992).

With the exception of the use of nondirective consultation as a format of partner-directed intervention (Sigafoos, Kerr et al., 1994), there has been little emphasis in partner instruction research regarding the appropriateness of the training formats selected for targeted partner groups. The literature on adult education suggests that the following format best supports the learning styles of adult learners and yields the most effective training impact: a combination of (a) workshops, (b) modeling, (c) practice in simulated and real settings, (d) feedback about performance, and (e) coaching during actual practice (Joyce & Showers, 1980; Korinek, Schmid, & McAdams, 1985; Sparks, 1983; Wade, 1984-
1985). Workshops are useful for presenting theories, skills, and strategies. Information and demonstration through modeling are needed to teach the basic content and convince participants of its importance. Practice is needed to develop fluency of the necessary skills. Guided practice in simulated or real settings will help to overcome the common problem of being unable to implement a strategy that appeared easy when first presented. And finally, one especially important component, essential for behavior change and maintenance, is feedback (Joyce & Showers; Sparks; Stevens & Driscoll, 1987).

The organization of the instructional strategy must also be suitable to the setting. Pyramidal training has been found an effective training strategy to train parents (Neef, 1995) and multiple groups of staff including group home supervisors (Parsons & Reid, 1995), direct care staff (Shore, Iwata, Vollmer, Lerman, & Zarcone, 1995), and teachers (McGimsey, 1995). In a typical pyramid, a cohort of staff is trained by “experts.” Once trained, this cohort trains a second cohort, and the second cohort trains a third cohort, and so forth. Pyramidal training is therefore extremely appealing for group home settings where staff turnover is usually high and it may not be efficient to have each new staff trained by an “expert” (Larson, Lakin, & Bruininks, 1998). To date, pyramidal training has not been used for training group home staff how to provide opportunities for communication to non-speaking children with developmental disabilities. The purpose of this study, therefore, was to determine the effectiveness of pyramidal training, using nondirective consultation combined with adult education formats, in instructing direct care staff to provide requesting opportunities to children with developmental disabilities and little or no functional speech.

**Method**

**Participants and Settings**

Participants included children with developmental disabilities residing in group homes and their assigned direct care staff partners. Three male children with developmental disabilities participated in the study. Children ranged in age from 7 to 10 years. One of the children, Steve, was diagnosed with pervasive developmental delay and autism, and Joe and Paul were diagnosed with developmental delay and autistic-like features. Information obtained by interviewing staff with the Communication Interview (Schuler, Peck, Willard, & Theimer, 1989), revealed that each of the children had no formal (i.e., symbolic) means of requesting. They tended to communicate their requests by touching items within reach or leading someone to a desired item. Even though the participants had been provided with communication displays their assigned direct care staff indicated that the participants did not use these displays to request preferred objects or activities. Each of the three children resided in a different group home, operated by the same private agency, where the study was implemented. The children functioned in the moderate to severe range of mental retardation in terms of IQ and adaptive behavior.

Seven adult direct care staff, who usually worked with the three children, participated in this study. These adult participants were selected because they were the most regular (albeit paid) communication partners of the participating children. Participants were all female and had worked in group home settings anywhere from a minimum of one year up to five years (mean = 3 years). They knew the particular child participant with whom they worked anywhere from one year up to four years and six months (mean = 2 years). All of the staff had at least a high school diploma with one of the staff holding an undergraduate college degree. None of the staff who served as communication partners had received formal instruction in AAC techniques prior to this study.

**Definitions of Dependent Measures**

Staff were observed to determine the number of requesting opportunities that each staff person provided using the missing-item, interrupted-chain, or delayed-assistance strategy. Operational definitions for counting a staff behavior as an opportunity were specific to each strategy and consistent with those developed by Sigafoos, Kerr et al. (1994). To be counted as an opportunity with the missing-item strategy, the staff not only had to engi-
neer the situation to create a missing item, but also had to wait at least 3 s before providing the missing item, as if waiting for the child to make a request. To be counted as an opportunity with the interrupted-chain strategy, the staff had to prevent the child from engaging, continuing, or completing an activity followed by at least a 3 s wait before allowing the activity to proceed. Finally, to be counted as an opportunity with the delayed-assistance strategy, the staff needed to approach a child who was clearly having difficulties completing a particular task and then wait for at least 3 s before providing the required assistance. It was not necessary for the child to make the request in order to be counted as an opportunity with any of the three strategies. When an opportunity was provided through any of the above strategies, the child’s response to the opportunity was recorded as an “unprompted request,” a “prompted request,” or “no response.” Operational definitions for prompted and unprompted requests were also taken from Sigafos et al. (1994). The response was considered unprompted when the child provided a response within 10 s of the staff approaching the child and before being provided some type of verbal, gestured, model, or physical prompt after the 3 s required delay. However, if the staff provided some type of verbal, gestured, model, or physical prompt after the required delay (i.e., 3 s), but before the child responded independently, then the child’s request was recorded as a prompted request.

The following dependent measures were used to evaluate the effectiveness of direct care staff instruction during baseline probes and intervention probes: (a) number of opportunities provided, (b) number of unprompted requests, and (d) number of prompted requests. Number of requesting opportunities provided was obtained by simply adding the number of opportunities provided using any of the three strategies. Number of unprompted requests and the number of prompted requests were obtained by adding the number of occasions when the child produced a communicative response that was accepted by the staff (as indicated by her reinforcement of that response).

Experimental Design

A multiple-probe design (Horner & Baer, 1978) across three cohorts of direct care staff was used to evaluate the effects of pyramidal staff instruction on the number of opportunities provided and the number of requests emitted by the children. The first cohort involved three staff-children dyads, including Staff A-Steve, Staff B-Joe, and Staff C-Paul. The second author, a certified Speech-language Pathologist, and the first author, who holds a Ph.D. in Special Education, trained this cohort of staff. Based on the nomination of the agency’s administration as the most stable and long-term employees, Cohort 1 staff was targeted to serve as trainers of staff in Cohorts 2 and 3. The second cohort included the same three children, although with different staff (Staff D, E, and F). The third cohort included only two of the three children (Steve, Joe) with yet another set of direct care staff (Staff G, and H). Paul was not included in the third cohort as the staff assigned to him took maternity leave following baseline.

Observation and Videotaping Procedures

Observations were conducted via a video camera to record the number and type of requesting opportunities provided and the number of prompted and unprompted requests. For each staff-child pair, 15 min observation sessions per activity were conducted approximately bi-weekly. The activities selected were recurring as part of the group home routine at approximately the same time on the days of data collection. For the first cohort, the activities for Steve, Joe, and Paul involved chair wiping, water play, and preparing lunch, respectively. For the second cohort, activities for Steve, Joe, and Paul included preparing lunch, mealtime, and bedtime routine, respectively. And the activities for the third cohort involved laundry for Steve and Chores/Table setting for Joe.

Sessions were videotaped by a research assistant (RA) with extensive experience in videotaping interactions. The RA positioned herself away from the staff and the child and remained as unobtrusive as possible to minimize observer effects. The video camera was mounted on a tripod and remained stationary.
during videotaping unless the staff and child moved out of the camera’s range, in which case adjustments were made to the camera’s position. Profiles of the staff and the child were made to allow a clear view of both.

Coding Procedures

Two research assistants (RAs) were trained to code the videotapes. One RA served as the primary rater while the second RA served as the independent observer to obtain interobserver agreement data (see Interobserver Agreement). The second RA was blind to the purpose of the study. Both observers were trained to follow the target behaviors (see above) prior to baseline by coding 15 min segments of two activities involving a staff-child pair that was not part of the research project. Both observers coded the videotapes in the presence of an author until a standard was established. Instruction continued until both observers achieved an agreement of at least 90% with the standard.

Procedure

Baseline probes. Baseline sessions were observed for one targeted activity for each staff-child dyad. The observer entered the room at the time when each activity was about to be conducted, located the staff and the child, and videotaped the activities. Each session lasted 15 minutes per activity. During baseline, the staff was informed that the observer was present to observe the children’s existing communicative behavior.

Pyramidal instruction. Between the last baseline session and the first intervention session, the staff from the first cohort participated in an instructional program, consisting of (a) the presentation of theory, skills, and strategies, (b) modeling, (c) practice of modeled skills in simulated environments, (d) generating others ways that the strategies could be used in targeted activities, (e) practice in real environments, (f) coaching during actual practice (i.e., steps c-e), and feedback about performance (i.e., during steps c-e). Components a through c were implemented as part of a one-day workshop involving first cohort staff and the authors.

The overview of theories, skills and strategies lasted approximately three hours and involved the following topics: the importance of the partner in promoting communication (Light et al., 1992), identifying communication environments, selecting vocabulary and symbols for different activities, strategies for creating opportunities for requesting (i.e., missing-item, interrupted-chain, and delayed assistance), and the techniques involved in using delayed prompting. As part of the presentation of strategies, the staff received a one-page description of the missing-item, interrupted chain, and delayed-assistance strategies developed by Sigafoos, Kerr et al. (1994). Following the presentation of strategies, the instructors modeled them. Modeling included the demonstration of each strategy for creating requesting opportunities along with delayed prompting through role-play using several hypothetical activities. Each strategy was modeled with each activity to instill that the strategies are flexible and not limited to only one particular activity. Staff was then asked to practice the strategies through role-playing using the same hypothetical examples while the instructor provided coaching and ongoing feedback.

In individual consultation sessions, the staff was then asked to generate ways these strategies could be used with their child in the activity targeted for intervention. Occasionally, instructors had to facilitate this process through guiding questions (e.g., how might the missing-item format be used with Steve?). Consultation sessions were also used for creating topic-specific communication displays for each activity. Using the ideas and displays generated, the staff practiced the use of the strategies in the targeted activity with their child in one 20-minute session. Instructors provided coaching and feedback throughout and following this session. Where appropriate, the list of ideas generated during consultation was revised utilizing this feedback.

Staff of the second and third cohort was trained by the first cohort staff using individual consultation sessions and a one 20 min practice session. In order to maintain efficiency, the staff from these cohorts did not partake in a workshop.

Intervention probes. Observations during intervention were identical to baseline sessions involving the target activity. Five minutes be-
fore each session, however, the observer and staff reviewed the one-page description of the three strategies and the accompanying list of ideas for using the strategies that had been generated through the consultation process and revised following practice in real environments. Using these ideas and strategy descriptions as a guide, staff was asked to provide as many opportunities for requesting as possible during the upcoming 15 min session. During the 15 min probe session staff was provided with no feedback because these probes constituted the basis for evaluating the effectiveness of intervention. After each session, however, the observer provided feedback to staff on the number and types of opportunities they had actually provided and the delayed prompting techniques used during the preceding session. Additional ways for using the strategies were then discussed.

Interobserver Agreement

Interobserver agreement checks were conducted by the second RA for dependent measures equally across baseline and intervention probes (25%) and across all staff-child pairs. After receiving instruction (see coding procedures), the RA recorded the target behaviors during all phases of the study. An agreement was scored when the primary rater and the second RA had recorded the same type of opportunity and the same type of request at the same clock time (to the nearest minute). Interobserver agreement was calculated by taking number of agreements divided by number of agreements plus disagreements and multiplying by 100%.

During baseline probes, interobserver agreement on opportunities provided across staff yielded a mean of 100%. Interobserver agreement on the type of requests yielded a mean of 100% as well. During intervention probes, interobserver agreement on opportunities provided across staff yielded a mean of 91% (range: 85–95%). Interobserver agreement on the type of request yielded a mean of 90% (range: 83–94%).

Results

Results for opportunities provided and requests made by the children are displayed in Figure 1. Because only a total of three requests by the children were unprompted, we opted to combine prompted and unprompted requests with the understanding that the majority of requests were prompted.

Baseline

During baselines, including extended baselines of Cohort 2 and 3, staff provided no opportunities for requesting through the missing-item strategy, interrupted chain strategy, or delayed-assistance strategy. In turn, no requests were observed by any of the children across cohorts. These data were consistent with the agency’s reason for seeking consultation services from the authors; that is, the lack of formal requesting among participating children.

Intervention Probes

Staff in all three cohorts displayed marked improvements in providing opportunities for requesting only after training for a specific cohort was initiated. At the same time, the children across cohorts increased their number of requests made; each cohort did so only once training was initiated. Although the level of requesting indicates that the children failed to make use of all opportunities provided to them, a direct relationship between opportunities provided and requests made is clearly indicated.

Discussion

The purpose of this study was to determine the effectiveness of pyramidal training, using nondirective consultation combined with adult education formats, in instructing direct care staff to provide requesting opportunities to children with developmental disabilities. Results clearly demonstrate that the intervention resulted in increased opportunities provided by staff along with concomitant increases in children’s requesting. Thus, pyramidal training appears to be a viable option to train direct care staff in promoting communication of non-speaking children with developmental disabilities.

Previous studies had indicated that children’s request behavior could be increased...
Figure 1. Number of opportunities provided and requests made. Open squares indicate opportunities provided and closed circles indicate requests made.
through the training of partners in providing more opportunities (Sack et al., 1992; Sigafoos, Kerr et al., 1994). This study adds to this literature in that it demonstrates a viable approach for “expert” trainers to train only a segment of direct care staff who in turn provide training to their peers without further direct involvement of the experts. The knowledge that it is effective to train the trainer may assist group home supervisors in ensuring that new staff, arising due to often rapid turnover, is readily trained by their peers (see Larson et al., 1998). This study also extends previous work on the effectiveness of pyramidal training with group home staff (Parsons & Reid, 1995; Shore et al., 1995) to a different group of clients whose primary needs relate to the development of communication skills. It should be kept in mind, however, that the pyramidal approach used in this study was somewhat different from previous applications. In typical applications, the first cohort trains the second cohort, and the second cohort trains the third cohort. In this study, the first cohort trained each of the subsequent two cohorts. This was more consistent with the expressed expectations and needs of the supervisory staff in the group home agency. Future research needs to be directed into the relative effectiveness of these various approaches to pyramidal training.

The vast majority of requests emitted by participating children were prompted rather than unprompted. There are a number of plausible explanations for this somewhat disappointing finding. First, delayed prompting was addressed more with the first cohort as they received the workshop in addition to individual consultations, practice sessions, and feedback. This may explain why the only three unprompted requests were demonstrated during interactions with the first cohort.

It is important to note that the staff-training package included a number of components (e.g., in-service training, written guidelines, modeling, feedback). It is unclear to what extent each of these components was necessary and contributed to the outcomes obtained. Similar results may have been obtained by using only the in-service training component, for example. While in-service training alone might represent a more efficient training approach, the absolute savings in terms training time would not seem to be so great so as to recommend exclusion of the other components. Still, it would require additional component analysis to determine an optimal training package. The present results do however suggest that the current package was effective and the pyramidal approach would seem to have made it an efficient way of training staff. The lack of treatment integrity data is a limitation. Future research would be improved by monitoring the extent to which the pyramidal training program is delivered as specified.

References


Houghton, R. D., Bronicki, G. J. B., & Guess, D.


Received: 10 March 2005
Initial Acceptance: 2 May 2005
Final Acceptance: 1 October 2005
Computer-Based Instruction for Purchasing Skills

Kevin M. Ayres, John Langone, Richard T. Boon, and Audrey Norman
The University of Georgia

Abstract: The purpose of this study was to investigate use of computers and video technologies to teach students to correctly make purchases in a community grocery store using the dollar plus purchasing strategy. Four middle school students diagnosed with intellectual disabilities participated in this study. A multiple probe across participants research design was used to evaluate the effectiveness of the treatment. Results indicated the program was effective at teaching the dollar plus purchasing strategy to three out of four participants and promoted generalization to the natural environment. Finally, limitations of the study, implications for practice, and future research questions are discussed.

Stokes and Baer (1977) highlighted several features of instructional programs that promote generalization. Two of those features relative to effective instruction of community skills (e.g. shopping, ordering from a menu, navigating public transit) are the programming common stimuli and training of sufficient exemplars. Programming the common stimuli generally requires being in the environment where those stimuli exist and programming sufficient exemplars requires time in that environment. This may become an obstacle to teachers planning community-based instruction [CBI] because of logistics and cost involved in providing instruction in natural environments (Wissick, Gardner, & Langone, 1999). Thus, teachers need to monitor efficiency of CBI to maximize the opportunities for their students and search ways to enhance their instruction. One possibility for this is to supplement CBI with classroom simulation. To expect generalization from classroom simulation to the natural environment, however, will require careful planning on the part of teachers to make sure that the simulations match the natural environment to the fullest extent possible.

To create realistic opportunities for students with disabilities to practice community-related skills in the classroom, teachers have to focus on making the simulation as close to real life as possible. Ideally, no teacher would want to teach a community skill solely in the classroom, but time in the community to practice skills sufficiently may be cost prohibitive and logistically difficult. If teachers are able to supplement their community-based instruction with classroom simulations, they may be able to stretch the utility of their overall instructional program (e. g. Mechling, 2004). To be useful though, the skills practiced or learned in the simulation need to generalize to the community setting.

Morse, Schuster, and Sandknop (1996) recommend focusing instruction on shopping skills in the broadest sense to incorporate the whole experience of going to the store. It may be possible to isolate certain skills that may be more easily simulated in the classroom than others and thus, leaving most of the instructional time available for skills that would be difficult to simulate in the classroom (e.g. asking for help when one cannot locate an item). Branham, Collins, Schuster, and Kleinert (1999) demonstrated that classroom simulation combined with in vivo instruction was more efficient for teaching check cashing, street crossing and letter opening than community-based instruction combined with video taped modeling or video taped modeling and classroom simulation.

There might be instances when teachers might choose to isolate certain skills to be taught primarily in the classroom and depend-
ing upon the quality of exemplars used in the simulations, the instruction could be effective in teaching certain skills. For example, Mechling and Gast (2003) provide an example of using a computer-based simulation to teach students to locate items in a grocery store using aisle signs as guides to locate items that were not specifically written on the aisle signs (e.g., if the aisle sign said Cake Mixes, the students learned that they would also find brownie mixes on that aisle). Similarly, Wissick, Lloyd, and Kinzie, (1992) combined video models with computer-based instruction to decrease the number of error’s students made on a shopping trip. However, there is limited research on how computer-based instruction can enhance other classroom based practices.

Therefore, the purpose of this study is to examine use of computers and video as an addition to an on-going classroom based intervention to teach money skills that had not previously fostered generalization to the community. The focus of this study is on using computer and video technology to teach students to pay correctly for a purchase in a store using the dollar plus or dollar more strategy (Colyer & Collins, 1996; Denny & Test, 1995; Schloss, Kobza, & Alper, 1997; Test, Howell, Burkhart, & Beroth, 1993). Prior to this study all of the students had learned to make purchases to whole dollar amounts (e.g., $4.00) with some combination of in vivo and classroom based instruction. The students were engaged in classroom-based tabletop simulation activities that were not facilitating generalization to the community. The computer-based program used in this study was designed to teach students to pay for uneven dollar amounts on the computer, and through the use of video models, was designed to facilitate generalization to the community. Specifically, this study is a systematic replication of Ayres and Langone (2002), where video based models and a computer interface was employed to teach students the dollar plus strategy to a group of students who had no previous experience with making purchases of any amount. The present study incorporated certain changes designed to increase the chances that this group of students would be able to generalize the skills learned on the computer to in vivo activities. Results of the original study published in 2002 indicated that changes in computer interface were warranted. These changes would allow the learners to have a more realistic visual presentation of the stimuli and provide them more specific feedback for incorrect responses.

Method

Participants

Four middle school students diagnosed with intellectual disabilities participated in this study (see Table 1). All students were 14 years of age and served in a self-contained classroom in a rural middle school with six other students. Participants were selected based on inclusion of IEP goals related to purchasing skills. All students in the class participated in community-based skills twice per week. The teacher emphasized vocational skills on one day and shopping or leisure skills on the other day. During the shopping exercises the previous year, students had learned to pay for items that totaled to whole dollar amounts.

Adam a young man with Down Syndrome, was able to accurately follow verbal directions. He could read elementary text and would attempt unknown words by phonetically sounding them out (recent test data were not available). He could also perform simple addition and subtraction. Most of his IEP goals were focused on independent living skills (e.g., food preparation, job skills, and domestic skills). Socially, Adam was very adept and friendly. He enjoyed participating in computer activities and had past experience working on literacy programs via the computer.

Emily also had Down Syndrome. Academically she could read some sight words and do simple addition problems. Like Adam, she was very social and worked well with classmates and teachers. She also had experience working on the computer and most of her IEP goals were related to functional living skills as well. Emily had a medical condition that occasionally resulted in an interruption of the sessions to provide her with medication and rest.

James participated in a previous study conducted by Ayres and Langone (2002) under the same pseudonym. In that study, James made little progress on the computer-based
instruction and failed to generalize sufficiently to the community activities. He was diagnosed with Down Syndrome and an unquantifiable hearing loss in one ear. Academically, James was working on basic sight word reading and basic number skills. He had experience working with computer-based instruction and was reported to enjoy playing video games on the computer. Socially, James was a friendly and polite student who enjoyed engaging in conversation with peers and adults.

Arnold exhibited strengths in both receptive and expressive communication. Arnold had difficulty speaking clearly. His IEP focused mainly on daily living skills with some functional academics. According to teacher reports, he had difficulty sitting still for long periods but enjoyed working on the computer.

**Settings and Materials**

This study took place in two settings. The first setting was a large national grocery chain. Students participating in the study used a grocery store lane that was staffed by a trained confederate (i.e., third author and an undergraduate student in special education) playing the role of the cashier. Individually, students took baskets of food to the cashier and made their purchases. The entire checkout line system (scanner, total on the computer screen, receipt printer etc.) functioned as if the student were making a genuine purchase from a store employee. The grocery items that students used for purchases were pre-selected to represent various dollar totals from $1.01 to $9.99. Students also used a stack of 12 one-dollar bills for their purchases that they placed in their wallets.

The second setting in which part of the study took place was the self-contained classroom at the students’ middle school. The classroom measured approximately 15m × 20m with 2/3 of the classroom set up like a house (a kitchen, laundry area, living rooms, bathroom), 1/3 as a traditional classroom with desks and computers. A computer station was partitioned from the rest of the class and

---

### TABLE 1

Psychometric Description of Participants

<table>
<thead>
<tr>
<th>Age</th>
<th>Stanford-Binet-IV b</th>
<th>Vineland Adaptive Behavior Scales b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Composite: 50</td>
<td>Composite: 58</td>
</tr>
<tr>
<td></td>
<td>Verbal Reasoning: 70</td>
<td>Communication: 47</td>
</tr>
<tr>
<td></td>
<td>Abstract Visual Reasoning: 48</td>
<td>Daily Living: 60</td>
</tr>
<tr>
<td></td>
<td>Quantitative Reasoning: 60</td>
<td>Socialization: 74</td>
</tr>
<tr>
<td></td>
<td>Short Term Memory: 50</td>
<td></td>
</tr>
<tr>
<td>Adam 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emily 14</td>
<td>Composite: 58</td>
<td>Composite: 68</td>
</tr>
<tr>
<td></td>
<td>(No subscales available)</td>
<td>Communication: 62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Daily Living: 63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Socialization: 86</td>
</tr>
<tr>
<td>James 14</td>
<td>Composite: 41</td>
<td>Composite: 58</td>
</tr>
<tr>
<td></td>
<td>Verbal Reasoning: 49</td>
<td>Communication: 59</td>
</tr>
<tr>
<td></td>
<td>Abstract Visual Reasoning: 52</td>
<td>Daily Living: 56</td>
</tr>
<tr>
<td></td>
<td>Quantitative Reasoning: 58</td>
<td>Socialization: 67</td>
</tr>
<tr>
<td></td>
<td>Short Term Memory: 45</td>
<td></td>
</tr>
<tr>
<td>Arnold 14</td>
<td>Composite: 38</td>
<td>Composite: 46</td>
</tr>
<tr>
<td></td>
<td>Verbal Reasoning: 49</td>
<td>Communication: 48</td>
</tr>
<tr>
<td></td>
<td>Quantitative Reasoning: 60</td>
<td>Socialization: 75</td>
</tr>
<tr>
<td></td>
<td>Short Term Memory: 44</td>
<td></td>
</tr>
</tbody>
</table>

*Stanford-Binet Intelligence Scale Fourth Edition (Thorndike, Hagen, & Sattler, 1986)*

*Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984)*
served as the location where students engaged in the computer-based probes as well as computer-based instruction. Students wore earphones while working on the computer to further reduce distractions for themselves and their classmates. In addition, students used a portion of the PROJECT SHOP CD-ROM (Langone, Clees, Rieber, & Matzko, 2003) specifically targeting purchasing skills (see Figure 1 for screen captures). PROJECT SHOP was a federally funded project focused on the development of a multimedia program to enhance and augment community-based instruction.

**Response Definitions and Data Collection**

The dependent variable was accuracy of response and was defined as a student beginning payment for an item within 5 s of the cashier announcing the total and completing the response by handing the cashier the correct amount of money (at least the stated amount but no more than $.99 over the stated amount) within 20 s of beginning the response chain. This response was scored as correct. If the student handed the cashier an incorrect dollar amount (too much or too little), or did not complete his or her response within 20 s of initiating payment, the response was scored as incorrect. If the student did not respond within 5 s of the cashier announcing the total, the response was recorded as a no-response error. Additionally, during community probes, the duration of student response was recorded. The timer began when the cashier announced the total and ended when the student turned his or her palm up to await their receipt and change.

During computer training, five different responses were recorded. An unprompted correct was scored when the student initiated a response within 5 s of the cashier announcing the total and accurately completed the response by handing the cashier the correct amount of money. The timer began when the cashier announced the total and ended when the student turned his or her palm up to await their receipt and change.

![Figure 1. Screen capture of computer program.](image-url)
response (i.e. paying at least the stated amount but no more than $.99 more than the stated amount) within 20 s of beginning the response. Differences in latency and duration requirements between the computer and in vivo were required to allow for reasonable interaction time with the computer interface. If the student began a response within 5 s and either did not pay the correct amount or took more than 20 s to complete a response, the trial was scored as an unprompted incorrect. If a student did not begin a response within 5 s and the computer prompted the student, the student had an additional 5 s to begin and complete a response. If the student completed the response accurately within 20 s of starting the response the trial was scored as a prompted correct. Similarly, if the student was prompted by the computer after not having responded for 5 s and he or she then initiated a response but failed to respond accurately or complete the response within 20 s, the response was scored as a prompted incorrect. Lastly, if the student did not initiate a response within 5 s of the computer delivering the S0 and then did not respond when the computer made a prompt, the trial was scored as a no-response.

Procedure

General procedures. Students in this study already used the dollar plus strategy to purchase grocery items in classroom simulations and it was determined (i.e., through analysis of baseline data taken for this study) that the intervention was not successful at helping students to generalize to the community. The first step in the study was to determine the student’s present level of competence with the strategy in the community. After students demonstrated a stable baseline on community-based probes, both for accuracy of response as well as duration of response, students took part in computer-based probes. The classroom instruction (table-top simulations such as the teacher saying “You owe me $4.55” and the students responding by counting out $5) continued so that from baseline to intervention only a single variable was changed (the addition of the computer-based instruction). The computer-based probes were designed to evaluate student baseline performance with the computer-training tool. After stable baseline performances were achieved the first student began intervention. Once a student responded correctly and without a prompt to 80% of the trials during the computer-based training, all students in their received another community based probe to assess generalization for the student who reached the computer criterion and to monitor maturation of the other participants. Then the remaining students completed another set of computer-based probes and the next student began intervention. For the first students who achieved mastery on the computer-based probe, community-based probes were continued following intervention to monitor maintenance of treatment effects. An additional probe was conducted at the conclusion of the study in a store at a local shopping mall where students had never previously made a purchase, this allowed for some limited estimation of generalization across environments.

Classroom based instruction. Classroom instruction took place prior to the beginning of this study. The first community probes represent the degree to which this intervention allowed generalization of the behavior from the classroom to the community. To reduce the number of variables altered from baseline to intervention, all students continued to take part in this classroom based instruction. Students sat around a kidney shaped table, each with a stack of 12 one-dollar bills. The teacher sat in the middle of the kidney shaped table with a small cash register, her data collection forms, a stopwatch and pen. To keep students engaged, the teacher proceeded in a round robin fashion giving each student one trial at a time. The order moved predictably from left to right.

The teacher began sessions by keying a total into the cash register and turning the cash register toward the first student. She then said to the student “Your total is ________.” After announcing the total, the teacher placed her hand on the table in front of the student to await payment. The student was allowed 5s to respond. If he or she responded correctly the teacher praised the student and gave them change and a receipt. If the student did not initiate a response within 5 s the teacher began to prompt the student. First the teacher would repeat the total and point to the total
on the cash register screen. If the student still did not respond, the teacher began to count aloud for the student to follow along while placing bills on the table. The teacher counted until she reached the dollar total (e.g. five for a total of $5.35), and then she said “and one more” to indicate that the student needed to place one more bill on the payment pile. Once a trial was complete the teacher moved to the next student; she continued around the table until all students had completed five trials.

Community-based probes. The purpose of the community based probes was to allow the students naturalistic opportunities to demonstrate their mastery of paying for grocery items, to monitor any generalization from the on-going classroom instruction, and to assess generalization of the behaviors demonstrated during computer-based instruction. Ideally probes would have been conducted over several days with the student making actual purchases at stores. To generate adequate data this would have been logistically impossible based on school system resources to provide community based instruction. Instead, similar to Ayres and Langone (2002), probes were staged in a community store. A confederate played the role of the cashier at the grocery store. Students were handed 12 one-dollar bills and told by the teacher to “Go to the cashier and pay for your groceries.” The cashier would ring up the groceries and announce the total following a randomized script to vary the verbal stimuli between possible presentations (e.g. “three dollars and thirty four cents” and “five twenty-five”).

After announcing the total the cashier stood with his or her hands at their sides and waited 10 s for the student to begin a response. If the student did not respond, the teacher called the student back to the end of the line and the next trial began. If the student did begin a response, the cashier either held his or her hand out to receive the money (if the student paid to their hand) or waited passively with their hands at their sides (if the student paid the money on to the counter). In either case, when the student turned his or her palm up to wait for their change, the trial ended and the cashier handed the student change and a receipt. The student walked to the teacher with the change and receipt and the teacher handed the student another 12 one-dollar bills and told the student to go back and try again. The student was not told whether they responded correctly or incorrectly but was verbally praised for “working hard.” Students completed five trials per probe session and at least three probes were conducted for initial baseline. When two sessions were conducted on the same day, each session was separated by at least 30 min. Participants in the study not engaged in probes were elsewhere in the store with school staff working on other IEP objectives.

Computer-based probes. These probes were conducted on the classroom computer. The teacher instructed the student to sit down at the computer and told them that they would be practicing purchasing items like they do in the store. The teacher reminded the student to use the dollar plus strategy. When the program began, students clicked on the video in the bottom of the screen to pay another dollar. Each time they clicked the video, the hands in the video moved another dollar to the counter. When students were finished paying they clicked the closed wallet finish button to the right of the purchasing video. The computer allowed the student 5 s to begin a response and 20 s to complete the response. The students did not receive any feedback from the computer or teacher during these probes. Students completed 10 trials during each probe session and each block of probes lasted at least three sessions across at least two days until data were stable.

Computer-based instruction. During computer-based instruction, [CBI] students sat alone at the computer to work on the program. Each session consisted of 10 trials and the students engaged in one session per day separated by at least 1 hr. The layout and presentation were identical to probes except that the computer provided feedback for student response. Students were allowed 5 s to begin a response after the cashier announced the total. If the student did not begin a response in that amount of time the computer repeated the S5 and the trial was scored as a non-response error. If the student did begin a response but failed to finish the response within 20 s or did not hit the “finished but-
ton,” this was scored as an unprompted error, the computer provided descriptive feedback (e.g. “remember to click the finished button”), and the next trial began. If the student began and finished a response and paid either too much or too little, the computer provided corrective feedback in the form of specifically telling the student whether they paid too much or too little, the computer then demonstrated the correct way to pay and the student was given another opportunity to make the payment (this was still scored as an unprompted error but allowed the student to practice the correct response after watching a computer model). If the student still did not respond correctly, the computer guided the student through the correct response by highlighting places on the screen for the student to click to respond correctly thereby making certain that the student moved through the response chain. If the student responded correctly a video played of the cashier thanking the customer and handing the customer change and a receipt. The next trial began in a similar fashion.

Modifications. After Adam reached criterion on the computer he did not immediately generalize the acquired behavior to the in vivo setting. Two learning trials (including prompts) that exactly mirrored CBI were provided prior to session 43, and he was reminded, “to pay just like on the computer.”

Inter-observer reliability and procedural reliability. Inter-observer agreement and procedural reliability data were collected for the dependent measures during at least 33% of community-based probes for each student. The classroom teacher (fourth author) who held a masters degree in special education and had experience in single subject research, acted as reliability observers. They stood approximately 1 m away from the primary data collector, 3 m from the cash register but within range to hear and observe all student and cashier actions. The percentage of inter-observer agreement was calculated by dividing the total number of agreements by the sum of agreements and disagreements and multiplying by 100. Procedural reliability data were collected by following a protocol checklist where, for each trial, the observer marked whether the cashier engaged in the correct behavior. The total number of correct behaviors was divided by the total number of steps in the protocol and multiplied by 100 to compute a percentage of procedural reliability. Procedural reliability was 100%.

The computer tracked all data during computer sessions therefore it was not necessary to assess inter-observer reliability. However, to assess procedural reliability during computer-based instruction the teacher did one probe and one instructional session per week during which no problems occurred. No procedural reliability data were collected on classroom-based training because these procedures had been going on prior to this study and no student performance data were gathered either during these sessions either. This was deemed unnecessary because the primary concern was with generalization of the skill. Baseline probes were used to demonstrate the level of generalization.

Experimental Design

A multiple probe across participants design (Tawney & Gast, 1984) was used to evaluate the effects of intervention. The first student began intervention with subsequent students beginning additional probes and then intervention following as the first student reached criterion (improvement of 50% or more over baseline). This continued until all students had received intervention. After students met criterion, all students received community probes again.

Results

Figure 2 shows student performance from baseline to intervention for Adam and Emily and Figure 3 shows performance for James and Arnold. Closed circles represent the percentage of correct responses during community probes and the open triangles represent student responses on computer-based probes. The first student to receive intervention, Adam performed poorly during baseline in vivo probes. After 12 sessions on the computer he began to answer 100% of the computer-based probe questions correctly but he did not generalize the behavior to the community (Sessions 22-24). Following the single two trial training sessions in the community, between session 42 and 43, Adam immediately began
responding correctly to all trials in the community and continued to respond at high accurate rates for the remainder of the study.

Emily exhibited variable performance during baseline in vivo probes reaching a high of 100% correct for one session. Without being able to stabilize her performance, the choice was made to begin intervention. Following intervention, Emily’s community performance decreased from the baseline highs. Her work on the computer was equally variable.

James did not answer correctly during any of the baseline in vivo probes. Once he began intervention, he slowly began answering problems correctly on the computer. At his first opportunity to demonstrate the behavior in the community following intervention (Session 101), James answered 80% of the probes correctly and improved to 100% on the next two community probes.

During baseline in vivo probes, Arnold showed low variable responding. Upon introduction of intervention, he began responding accurately on the computer and this performance quickly generalized to the in vivo setting where he accurately responded to 60%, 40% and 60% of the probe questions correctly in his final community session.

Discussion

Based on visual analysis of the data, the program was effective at teaching the dollar plus purchasing strategy to three out of four participants and promoted generalization to the natural environment. The classroom teacher reported that the fourth participant, Emily, had difficulty controlling her medical condition during the study and the results appeared to be causing problems with her ability to concentrate in the classroom and on community skills. Normally, when this student is able to control her condition her attention and concentration are adequate to allow her to perform well in school. The overall impact of this program demonstrated positive effects for the other three students.

The remainder of this discussion considers two primary things. The first thing considered are the differences between the results from...
the Ayres and Langone (2002) study and the current investigation including some cautious suggestions about how these differences may have influenced the outcomes. Second, a more global discussion of how this study fits with the extant literature on community-based instruction and computer-based instruction and the directions the current results suggest for further research.

This study differed from Ayres and Langone (2002) in several ways that may have influenced the outcome. First, students participating in Ayres and Langone were significantly younger than those participating in the current study (mean age of participants in 2002 was 6.3 where as the mean age of the current study was 14). As might be indicated by their age, students in the present study also had significantly more experience in community-based instruction and this could have influenced their performance. In addition, during the present study, students received concurrent instruction in the classroom. Although these data show no evidence of this practice in the classroom directly improving student performance (note that student performance in the community did not improve until after computer-based instruction), it is possible that the additional practice in community environments may have influenced the outcomes.

A second difference from the Ayres and Langone (2002) study revolved around the instructional design of the computer-based instruction. In the 2002 study, the computer-based instruction taught the purchasing skills with a constant time delay procedure [CTD] (Wolery, Ault, & Doyle, 1991) with a video model as the controlling prompt and the computer provided general feedback for incorrect responses (e.g. “this amount is not right”) before moving to the next trial. This does not mean to imply that CTD is an ineffective procedure to use in computer-based instruction, sufficient literature exists to support its use (e.g. Mechling & Gast, 2003). In the current study however, the computer did not use a
CTD procedure; rather if students made any error, the computer provided specific feedback (e.g. “You paid too much/too little). Then, the computer program provided a model of the correct response. If students still failed to answer correctly the computer guided the student through the response process by prompting with highlighted targets on the screen where the student would need to click. This made sure that students had to respond correctly before continuing to the next trial. By forcing students to make a correct response on each trial, the computer may have allowed the student the greater opportunity to learn the topography of the correct behavior.

A third design difference that separates this study from Ayres and Langone (2002) is the user interface. In that study, students saw a line of one-dollar bills across the bottom of the screen that they needed to click on to respond. Once a dollar had been clicked, the bill moved to the counter and was “used.” In the current study, the interface the students saw depicted a hand holding a stack of dollar bills over the counter. As the student clicked the hand, they saw video of the hand putting the bill on the counter. This perspective looked more like what a typical customer sees as he or she hands money to the cashier and may therefore, have possibly influenced generalization. We see this feature as a significant improvement over the interface created for the original study.

While the quality of computer-based instruction can and should be judged by the impact it has on a student’s behavior, designing effective programs requires isolating those features that are the most powerful and useful for students in learning. This program combined video models, and simple feedback to improve student performance. Further evaluation of other components though (e.g. pace of the models), complexity of feedback, explicit use of an errorless learning procedure (e.g. constant time delay) would provide software designers with more information to develop more powerful software. In addition, during the course of this study students were practicing the skills outside of the computer program. The degree to which genuine practice and simulated practice influence student outcomes requires further exploration. The integration of various modes of instruction may impact how rapidly students acquire the targeted skills.

In the final analysis, this study adds to the small, but growing literature that demonstrates that computer-based video models can be used effectively either with or without other simulations and community-based instruction to effectively teach functional skills. It appears that such instruction that uses technology can work with community-based activities to improve the efficiency of the instruction and possibly save time and other resources. Further research is needed to determine the best combination of video models with in vivo activities and also whether first person or third person models work best.

References


Received: 10 March 2005
Initial Acceptance: 7 May 2005
Final Acceptance: 15 September 2005
Increasing the Skills of Children with Developmental Disabilities through Staff Training in Behavioral Teaching Techniques

Anna-Lind Pétursdóttir and Zuilma Gabriela Sigurdardóttir
University of Iceland

Abstract: Two staff members working in different preschools received training in the form of instructions, modeling, immediate feedback, and social reinforcement of good performance. Effects of training were assessed with a multiple baseline ABCDCDCD design. Correct use of basic behavioral teaching techniques increased from 16–31% to 92–95% and the rate of learn units increased from 0–0.03 to 2.3–2.4 per minute. A multiple baseline design revealed skill improvements in two 2-year-old boys with developmental disabilities. Skills of instructors and children generalized across teaching settings. Follow-up measures showed maintenance and additional generalization of skills. The results confirm the positive effects of increased rate of learn units on teaching effectiveness and students’ skills, and the importance of high achievement criteria for generalization.

Effectiveness of teaching with discrete trials has been demonstrated in numerous studies (e.g., Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovaas, 1987; Matson, Benavidez, Comptom, Paclawskyj, & Baglio, 1996). A discrete trial is comprised of a discriminative stimulus, a response by the student, immediate delivery of consequences for the response, and a short intertrial interval (Koegel, Russo, & Rincove, 1977). Several techniques, such as prompting, fading, task analysis, and shaping are used with discrete trial teaching (Anderson, Taras, & Cannon, 1996; Leaf & McEachin, 1999) and it is most often used in a one-to-one format although group teaching arrangements can be successful as well (Taubman et al., 2001).

Discrete trial teaching has several advantages (Leaf & McEachin, 1999; Sundberg & Partington, 1998), one of the most important being the high rate of learn units it makes possible. A learn unit consists of a three-term contingency comprised of antecedents, responses, and consequences that interlock with each other during teacher-student interaction (Catania, 1998; Greer, 1994). Research has shown a positive correlation between the rate of correctly performed learn units by a teacher and student achievement (Greer, McCorkle, & Williams, 1989) as well as direct functional effects of the rate of learn units on students’ performances (Albers & Greer, 1991; Ingham & Greer, 1992) and student objectives met (Selinske, Greer, & Lodhi, 1991).

Learn units can be observed in any educational setting, but typically, the rate of learn units in traditional teaching is very low (Albers & Greer, 1991; Greer, 1994; Ingham & Greer, 1992). Thus, a considerable gap between research on effectiveness of teaching methods and common practices in regular schools seems to exist (Jahr, 1998). Research has shown positive effects of behavioral staff training on the rate of learn units and teaching effectiveness (Albers & Greer; Ingham & Greer), however, such techniques have seldom been used with preschool personnel (Crow & Snyder, 1998). In those studies where
behavioral staff training has been applied in a preschool setting, positive results have been reported regarding teachers’ use of the trained techniques (Peck, Killen, & Baumgart, 1989) and the rate of learn units (Selinske et al., 1991).

Nevertheless, even comprehensive training that results in significant improvements in staff performance does not guarantee positive effects on the skills of clients. In a study by Smith, Parker, Taubman, and Lovaas (1992), for instance, an intensive, 1-week workshop in behavioral theory and treatment techniques, which resulted in increased treatment skills at the workshop site, did not have any effect on group-home client functioning. Hence, it is essential to incorporate client data in the evaluation of staff training procedures (Harchik, Sherman, Hopkins, Strouse, & Sheldon, 1989; Jahr). Hitherto, few studies have included clients’ skills as a dependent measure of the effects of staff training (Demchak, 1987; Jahr, 1998), although the number of studies has been increasing (e.g., Schepis, Reid, Ownbey, & Parsons, 2001). However, research is lacking on effects of behavioral training for preschool staff on the skills of children with developmental disabilities (Crow & Snyder, 1998) and on generalization of teaching skills across settings and tasks (Jahr).

Discrete trial teaching of children with developmental disabilities usually starts out in a one-to-one setting before group teaching settings are considered (Lovaas, 1996; Handleman, Harris, Cristoff, Fuentes, & Alessandri, 1991). Thus, training of instructors often takes place in a one-to-one setting. Although generalization of teaching skills to new settings is essential to enhance learning of students in a variety of settings, research has rarely assessed generalization of skills acquired during staff training (Jahr, 1998). Only a small number of studies have included measures of generalization of teaching skills across teaching tasks (Koegel et al., 1977; Mörch & Eikeseth, 1992; Kissel, Whitman, & Reid, 1983; Thorisdottir, 1993). Several factors can influence generalization (Stokes & Baer, 1977). One important variable that affects generalization is training with sufficient exemplars until high rates of behavior are displayed. High mastery criteria have also been found to increase the likelihood of generalization of skills across tasks and time (Arco & Millett, 1996; Koegel et al.; Parsons, Reid, & Green, 1993).

The aim of this study was to: a) assess effects of staff training in behavioral techniques on instructors’ skills in the use of discrete trials and rate of learn units during their teaching, b) measure effects of changes in the staff’s teaching on the skills of children with developmental disabilities, c) evaluate generalization of instructors’ and children’s skills from a one-to-one teaching setting to a group setting and, d) assess generalization of the staff’s acquired skills to new teaching tasks.

Method

Participants

Instructors. Two instructors, each from a different preschool, who were specifically allocated to attend to each child participant during most of the children’s attendance time, volunteered. Dora was a 35 year-old paraprofessional with an elementary school education and 5 years of work experience at the preschool. She had been attending to David for 5 months before the beginning of this study, providing general assistance in various activities in the classroom. Dora had some basic knowledge of behavior management but no training or experience in applying behavioral teaching techniques such as discrete trials. Hanna was a 52 year-old special education preschool teacher and assistant director of the preschool. She had 33 years work experience with young children, but no training in applying behavior principles to teaching. The children in the study were the first children with developmental delays that the instructors worked with.

Children. Two children with developmental disabilities participated in the study. David was a 25 month old boy with Down’s syndrome and a developmental index of 55 according to Bayley Scales of Infant Development—Revised (BSID-II, Bayley, 1993). David was non-verbal, but used gestures and a few simple signs to communicate. He had some imitation skills and limited verbal comprehension. Adam was a 27 month old boy with developmental delays of unknown origin and a developmental index of 50 (BSID-II, Bayley). Adam was non-verbal and did not show any clear signs of
comprehending spoken language or symbols. He possessed no imitation skills and was not able to participate in classroom activities without manual guidance. Neither of the children had been exposed to discrete trial teaching before the study.

Setting

The study took place in two public preschools in Reykjavík, Iceland, which had facilities for one-to-one teaching. Preschool 1 served 56 children in three units. In David’s unit there were 17 children ages 1 to 3 years, including three with developmental delays. One-to-one teaching took place in separate rooms adjacent to the main activity room of the unit. During teaching sessions, the participants sat on child-sized chairs or pillows on the floor, facing each other. To their side was a table for materials. Group instruction took place in one of the main activity rooms of the unit, with the child seated alongside two to three other children facing the instructor.

Preschool 2 served 54 children in three units. In Adam’s unit, there were 14 typically developing children, aged 1 to 3 years. One-to-one teaching usually took place in a special education room, which was separate from the unit. Participants either sat on pillows on the floor or at a regular sized table with the child seated in a high chair sideways to the instructor. Group instruction took place in the main room of the unit or in the assembly hall (gym) with the participants sitting at a table or on pillows on the floor. In the group instruction sessions, two to three other children were present but engaged in other tasks.

Dependent Variables

Instructors. The target behaviors of instructors were: presentation of instructions, use of prompts, delivery of consequences, and the rate of complete learn units (frequency per minute). Definitions of target behaviors were drawn from previous teacher training research (Arco & Millett, 1996; Greer & McDonough, 1999; Koegel et al., 1977). Correct presentation of instructions was defined as a request directed to the child to perform a physical action, stated in a clear, specific, and consistent manner. Correct prompting was defined as any additional assistance (physical guidance or other) provided within 2 seconds from the presentation of an instruction, enabling the child to perform the physical action entailed in the instruction. Consequences were defined as the instructor’s reaction to the child’s response to an instruction. Consequences were considered correct if they were delivered immediately and contingent on the child’s response (i.e., distinctively positive following a correct response or a neutral “no” following an incorrect response). A learn unit was scored as correct if the use of instructions and consequences met the definitions for correct application. A more detailed description of the target behaviors can be obtained from the first author upon request.

Children. Four skills from the children’s Individual Education Plans (IEPs) were selected for intervention: imitation of object use, imitation of gross motor movements, following instructions, and receptive labeling of objects. Correct imitation was defined as performing the same action as the instructor. Following verbal instructions correctly was defined as performing the action the instruction entailed. Correct receptive labeling was defined as touching the object mentioned in the instruction. In all cases, responding had to occur within 5 seconds to be scored as correct and be free of errors (i.e., self-corrections were not scored as correct). Both children had previously been exposed to tasks involving these target behaviors but had difficulty acquiring them.

Dependent Measures

Teaching skills. Teaching sessions were videotaped by placing recording equipment on a tripod, 2-5 meters away and side-on to the participants. Instructors’ target behaviors were assessed by analyzing a total of 35 5-minute segments of Dora’s teaching sessions and a total of 43 5-minute segments from Hanna.

The three term contingency, or learn unit, was the unit of observation during these segments. The beginning of a unit was marked by delivery of a discriminative stimulus (S0) and ended with delivery of consequences. Each instructor’s target behavior was scored as either correct or incorrect, and prompting was
also recorded by type. Measures of instructors’ target behaviors were calculated by dividing correct use of each component of a discrete trial with the sum of the correct and incorrect use of that component and multiplying the result with 100 to get a percentage correct. Rate of learn units was calculated by dividing the number of correctly executed units in each 5-minute segment by 5.

Children’s skills. Each instructor scored the child’s response in between the teaching trials or during teaching breaks. Measures of child behavior were calculated by dividing number of correct responses by the total number of responses (i.e., the sum of correct, prompted, and incorrect responses) and multiplying the result with 100. The number of trials needed to teach each item within skill domains was also summarized.

Performance Criteria

Instructor’s performance criterion was an average of at least 80% correct teaching across target behaviors in three consecutive sessions. Child’s performance criterion was at least 80% correct responding in three consecutive teaching sessions. When the child had learned at least six items in a skill domain and both instructor and child had reached their performance criteria, the training was moved from a one-to-one teaching setting to a group setting. Simultaneously, a new task was introduced for teaching in the one-to-one setting. Thus, experimental phases overlap with this regard.

Experimental Design

Two experiments with an ABCDCDCD with follow-up within-subject design were conducted simultaneously with multiple baselines across instructors. Dora and David participated in the former experiment, Hanna and Adam in the latter. Follow-up measurements and probes on generalization of teaching skills to new tasks took place 1 and 4 months after staff training ended. Effects of changes in the staff’s teaching techniques on the children’s skills and generalization were assessed with a multiple baseline across skill domains.

Experimental Conditions

Pretests. Prior to baseline a preliminary assessment of the children’s current functioning was conducted. The first author assessed each child’s ability in several skill domains, such as imitation skills, instruction following, and receptive and expressive labeling. This assessment was done to facilitate the selection of skill domains to target in the study.

Baseline (A). Baseline observations of instructor’s teaching performance took place during daily one-to-one instruction sessions. Five-minute videotaped segments of teaching were collected and analyzed for each instructor until a stable baseline was reached. Assessment of the children’s skills took place in a one-to-one setting and was recorded for evaluation of inter-observer agreement.

Workshop (B). The first author held a 5-hour workshop for each instructor, each child’s parent and one other preschool staff person. The workshop consisted of 3 hours of lectures on basic principles of applied behavior analysis accompanied by written handout (eight pages) and instructions on discrete trial teaching (three pages) as well as video clips of discrete trial teaching (total of 10 minutes) and discussion. The remaining 2 hours consisted of modeling discrete trial teaching with the target child, rehearsal of discrete trials with prompts, and performance feedback administered by the first author. On the two days following the workshop, instructors videotaped two 15-minute discrete trial teaching sessions under one-to-one conditions. Three 5-minute segments of their recordings were analyzed.

Training in one-to-one setting (C). Instructors were trained in the one-to-one instruction resource rooms, one to four times per week, for 6 weeks over a 10-week period. Each training session lasted 15-60 minutes and entailed the same procedures applied in the latter part of the workshop (except the video clips), delivery of prompts, occasional modeling, and performance feedback referencing the material covered in the workshop. Number and length of training sessions decreased as the instructors’ skills increased.

Generalization to group setting (D). Participant instructors were asked to continue teaching a skill in the same way as before with two
to three other children present. No further directions, prompts or feedback were provided. Taped segments of three to four teaching sessions were then analyzed.

Children’s responding was assessed in the same way as before. The instructor scored whether the response was correct, incorrect or prompted and the proportion of correct responding within each skill domain was calculated.

Follow-up measures. Follow-up measures were taken 1 and 4 months after training ended. The instructors’ teaching accuracy and children’s skills were measured in the same way as before. Instructors were simply asked to teach the children as usual, no instructions or prompts were provided. Sessions were taped as before.

Generalization of teaching skills to teaching new tasks. Instructors’ ability to teach a new task with written directions only was assessed 1 and 4 months after training ended. The first 5 minutes of each teaching session were recorded for evaluation.

Procedure

The study lasted approximately 3 months. After 3 to 5 days of baseline measures, the instructors participated in separate 5-hour workshops, spread over 2 days, followed by a 2-day evaluation period of the effects of the workshop. Dora’s training comprised 23 sessions that spread over 10 weeks. The training period was interrupted due to sick leaves and a 2-week vacation period. During the remaining 6 weeks, instruction of David took place on 34 days, 90-120 minutes per day. Hanna’s training comprised 20 sessions that spread over an 8-week period, which was interrupted by sick leaves and a 3 days vacation. On the remaining 6 weeks, Adam was instructed on 31 days, 60-140 minutes per day.

Instruction sessions usually took place twice a day, 30-110 minutes at a time. Each session was divided into 2–15 minute long teaching intervals, which included 5 to 50 discrete trials each. In between teaching intervals, contingent on a correct response, the child was allowed a 2–8 minute free-play break. Teaching focused on one item within a task, until a mastery criterion was achieved. A variety of reinforcers were used for correct responding, including listening to songs/singing, opportunity to play with a favorite toy, and having bubbles blown. On rare occasions edible reinforcers, such as raisins or cookie bites were provided.

Observer training. An undergraduate psychology student was trained to assist the first author in observing and scoring both instructor and child behavior. Training consisted of studying scoring instructions and practicing scoring participants’ behaviors from videotapes. After about 5 hours of training a criterion of 85% or higher occurrence agreement across observers was attained.

Inter-observer agreement. Seventeen to 50% of observations of instructor target behaviors in each experimental phase were checked for agreement, amounting to 39% of total observations. Observers used videotape recordings from teaching sessions to measure inter-observer agreement on observations of instructors’ performances. One three term contingency unit was watched at a time and the video paused while the performance was scored. Occasionally one of the observers asked for the trial to be shown again but did not enclose the reason for it. After independent scoring of the segments, occurrence agreements were calculated for each target behavior. Percent agreements were calculated by dividing the sum of agreed occurrences by the sum of agreed and disagreed occurrences and multiplying by 100. A summary of inter-observer agreement on observations of instructor performance is presented in Table 1. On average, agreements for all behaviors were above 95%.

Inter-observer agreement on children’s performance was measured in 24 to 72% of responses in each skill domain, a total of 39% of

<table>
<thead>
<tr>
<th>Instructor Target Behavior</th>
<th>Range</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation of S’s</td>
<td>85–100%</td>
<td>97%</td>
</tr>
<tr>
<td>Use of prompts</td>
<td>80–100%</td>
<td>95%</td>
</tr>
<tr>
<td>Delivery of consequences</td>
<td>87–100%</td>
<td>97%</td>
</tr>
<tr>
<td>Total average</td>
<td></td>
<td>96%</td>
</tr>
</tbody>
</table>
all measurements. The observers watched the teaching sessions and scored the child’s behavior simultaneously, but independently. At the end of the teaching session the instructor’s and the observer’s scoring of the child’s behavior were compared and agreement assessed. Percent agreement was calculated as described earlier. A summary of inter-observer agreement measures of child behaviors is presented in Table 2. On average, agreements for all behaviors recorded were above 90%.

Social Validity

After the experiment ended, instructors were asked to answer a 20-item questionnaire on the perceived usefulness of several components of the training and the effects they felt each component had on their skills and the children’s skills. A 5-point Likert scale was used with each question, with 1 standing for Not useful at all and 5 for Very useful.

Results

The staff training procedures had a clear effect on teaching skills of the instructors. Accuracy of their teaching improved and the rate of learn units increased. Moreover, independent correct responding of the children increased in all skill domains with one exception. There also was an indication of accelerated learning rate for Adam. Social validity measures revealed general satisfaction with the training procedures on behalf of the instructors and a perceived usefulness of them with regard to increasing their teaching skills and the skills of the children.

Changes in teaching. The workshop had a clear effect on the teaching techniques of both instructors. Overall, the application of three-term contingencies, or discrete trials, increased from 16–31% correct on average during baseline to 92–95% correct on average during intervention phases and learn units increased from 0.03 to 2.3–2.4 per minute on average for both instructors.

Figure 1 demonstrates to what extent teaching of the instructors was in accordance with the discrete trial teaching criteria. During baseline, Dora presented 6% of S^D’s correctly on average, 55% of prompts, and 16% of consequences. After the workshop, there was a clear increase in level of teaching accuracy. While teaching gross motor imitation, Dora presented on average 64% of S^D’s correctly, 91% of prompting, and 83% of consequences. Consultation in the one-to-one setting led to a further increase in teaching accuracy; resulting in an average 87% of S^D’s used correctly, 98% of prompting, and 95% of consequences.

Moreover, Dora’s skills in using discrete trials to teach gross motor imitation generalized to group settings without further training. While teaching gross motor imitation in a group setting the use of S^D’s was 98% correct on average, prompting was 100% correct on average and delivery of consequences 89% correct on average (see Figure 1).

In the second consultation phase, while teaching following instructions, Dora presented 91% of S^D’s correctly on average, 98% of prompts, and 94% of consequences. In the group setting, 96% of S^D’s were correctly presented on average, 100% of prompts, and 85% of consequences.

In the third task targeted during staff training, receptive labeling of objects, Dora presented on average 82% of S^D’s correctly, 98% of prompts, and 94% of consequences. Follow-up measures at 1 and 4 months after training ended showed that Dora maintained high teaching accuracy while teaching tasks targeted during staff training. The use of S^D’s was 100% correct, prompts were 94–100% correct, and consequences 92–94% correct.

Measures of generalization of teaching skills to new tasks showed that Dora was able to teach new tasks by following only written di-

| TABLE 2 |
|------------------|------------------|
| **Percentage of Inter-Observer Agreement Ranges and Means For Child Behaviors** |
| **Child Target Behaviors** | **Range** | **Agreement** | **Means** |
| Gross motor imitation | 83–100% | 92% |
| Imitation of object use | 82–100% | 93% |
| Verbal instructions | 78–100% | 90% |
| Receptive labeling of objects | 83–100% | 90% |
| Total average | | 91% |
rections. While teaching matching of objects 1 month after conclusion of staff training, Dora presented 95% of SDs correctly, 90% of prompts, and 81% of consequences. While teaching receptive labeling of body parts 4 months post training, Dora used 96% of SDs correctly, 67% of prompts, and 79% of consequences correctly as shown in Figure 1.

Hanna showed a similar increase in discrete trial teaching skills. Figure 1 shows that during baseline, on average 4% of SDs were correctly presented, 29% of prompts, and 15% of consequences. The workshop led to an increase in level of teaching accuracy as well as an upward trend. While teaching imitation of object use, Hanna presented on average 36% of SDs correctly, 77% of prompts, and 46% of consequences. Consultation in the one-to-one setting led to a further increase in level of teaching accuracy; on average 96% of SDs were correctly presented, 93% of prompts, and 89% of consequences. In addition, variability decreased considerably.

Hanna’s skills in using discrete trials to teach imitation of object use generalized to group settings, with every aspect of the teaching (SDs, prompting, and consequences) 100% correctly performed (see Figure 1).

While teaching to follow verbal instructions, Hanna presented 99% of SDs correctly, 99% of prompts, and 98% of consequences. The teaching skills generalized without further training to a group setting where every SD was correctly presented, 93% of prompts on average, and 93% of consequences as shown in Figure 1.

In the third task targeted during training, gross motor imitation, Hanna presented on average 99% of SDs correctly, 96% of prompts, and 100% of consequences (see Figure 1). The teaching skills generalized to the group setting where 98% of SDs were correctly presented on average, 100% of prompting, and 97% of consequences.

As can be seen on Figure 1, effects of training to teach the first task transferred to teaching the remaining tasks. Hanna immediately showed high teaching accuracy when teaching to follow verbal instructions and imitation of object use. The performance criteria for teaching in group settings were also attained in gradually less time across tasks. This reflects both Hanna’s improved teaching skills and Adam’s accelerated learning rate.

Follow-up measures at 1 and 4 months after training ended showed that Hanna main-
tained high teaching accuracy while teaching tasks targeted during staff training. The use of SP's was 92–100% correct, prompts were 100% correct, and consequences 92% correct.

Measures of generalization of teaching skills to new tasks showed that Hanna was able to teach new tasks by only following written directions. One month after staff training ended, every SP and prompt was used correctly and 88% of consequences, while teaching matching of objects. During teaching receptive labeling of objects 4 months post training, Hanna used 94% of SPs correctly, 100% of prompts, and 89% of consequences (see Figure 1).

Changes in prompting. The intervention led to a decrease in the use of prompts by both instructors (see Table 3). Thus, the proportion of independent responding on behalf of the children increased greatly during the intervention phases.

Increases in learn units. The average rate of learn units increased from 0–0.03 per minute during baseline probes to 2.3–2.4 per minute during the experimental phases.

As can be seen in Figure 2, there were no learn units in Dora’s teaching sessions during baseline, due to incorrect presentation of SPs and/or incorrect delivery of consequences. After the workshop, there was a clear upward trend in the rate of learn units as well as an increase in level reflected in an average of 1–2 learn units per minute while Dora taught gross motor imitation. Further increase in rate of learn units occurred after consultation took place, resulting in 3.2 units per minute on average. When teaching gross motor imitation was moved to the group setting the rate level dropped to 1.4 units per minute on average. Thus teaching in group settings included only half of the learn units compared to teaching in a one-to-one setting. This considerable drop is probably due to less time allocated to teach David while simultaneously attending to the other children.

While teaching the second task, following verbal instructions, Dora used 2.6 learn units per minute on average. Figure 2 shows an upward trend in the rate and no drop in level when the teaching was moved to a group setting, with Hanna maintaining an average rate of 2.3 learn units per minute. During teaching of receptive labeling, 2.1 learn units occurred on average per minute. Due to time limitations, teaching of the third task was not transferred to a group setting.

During baseline, only one learn unit occurred in Hanna’s observed teaching sessions (see Figure 2), which amounts to an average rate of 0.03 learn units per minute. After the workshop, there was an upward trend in rate of learn units, as well as increase in level, reflected in an average of 1 learn unit per minute when teaching imitation of object use. During consultation the average rate of learn units became 1.7 learn units per minute. This rate of learn units was maintained when teaching was transferred to a group setting, and even increased considerably during the last observation, leading to an increase in average to 2.6 learn units per minute.

While teaching the second task, following verbal instructions, Hanna used 2.3 learn units per minute on average and 2.6 learn units per minute on average in the group setting. The rate of learn units again showed an upward trend in the last staff training task, gross motor imitation. There, Hanna used 3.2 learn units per minute on average and the rate remained the same during teaching in the group setting, albeit with greater variability.

Improvements in children’s skills. Improvements in the children’s skills were considerable in three out of four tasks targeted in staff training. Figure 3 shows correct responding of David in skill domains targeted in staff training. David imitated 25% of gross motor movements on average during baseline. After Dora had attended the workshop, the level of correct imitation of David increased, reflected in an average of 69% correct responding, and during training in the one-to-one setting the

<table>
<thead>
<tr>
<th>Type of Prompt</th>
<th>Baseline</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dora</td>
<td>Hanna</td>
</tr>
<tr>
<td>Physical prompts</td>
<td>33%</td>
<td>29%</td>
</tr>
<tr>
<td>Other prompts</td>
<td>43%</td>
<td>30%</td>
</tr>
<tr>
<td>No prompts</td>
<td>38%</td>
<td>54%</td>
</tr>
</tbody>
</table>
Figure 2. Rate of learn units during baseline (A), following workshop (B), during one-to-one instruction (C), in group settings (D), at follow-up (FU), and when teaching new skills (Gen). Note: Phases separated by a dotted line show data that overlap in time. 1) Teaching in receptive labeling was not conducted in a group setting due to time limitations.

Figure 3. David’s performance during baseline and after his instructor had attended workshop (B), during one-to-one instruction (C), in group settings (D), and at follow-up (FU). 1) Teaching in receptive labeling was not completed due to time limitations.
overall level remained similar, with an average of 71% correct responding. Imitation skills of David generalized to the group setting, resulting in 87% correct responding. Correct responding to verbal instructions during baseline was 13% on average but after staff training began, there was a clear increase in level, to an average of 65% correct. Also, skills generalized to the group setting, resulting in 91% correct responding on average. During baseline, the performance of David in receptive labeling was 8% correct and correct responding increased to 48% correct during staff training. However, since the task only involved two choices, the performance was only at chance level. Several prompting approaches proved unsuccessful in teaching David receptive labeling of objects and due to time limitations teaching had to be terminated without clear positive results.

Figure 4 shows Adam’s correct responding in skill domains targeted during staff training. Adam did not imitate any object use during baseline, but after Hanna had attended the workshop, Adam’s correct responding showed a clear increase in level, to 36% on average. During staff training there was an upward trend in correct imitation of object use, with the average performance improving to 64% and generalizing to the group setting, resulting in 74% correct responding on average. Correct responding to verbal instructions during baseline was 3% on average but after staff training began, it increased to 73% correct on average and generalized to the group setting, resulting in 79% correct responding on average. Adam’s increased skills in imitation of object use did not result in increased skills in gross motor imitation. Correct responding in the latter skill domain remained low throughout baseline, 6% on average. However, when gross motor imitation was targeted in training, correct responding showed an upward trend as well as an increase in level, reflected in an average of 49% correct. The gross motor skills were maintained in the group setting, with an average of 98% correct responding.

Figure 5 shows changes in amount of trials needed to teach David several items within skill domains targeted in training. The

![Graphs showing Adam's performance](image)
amount of trials needed did not decrease as more items were taught; no clear pattern of acceleration did emerge in the learning of David.

Figure 5. Trials needed to teach David items within two skill domains.

Figure 6 shows the amount of trials needed to teach Adam several items within skill domains targeted in training. In general, trials needed to teach each item within the skill

Figure 6. Trials needed to teach Adam items within three skill domains.
domains decreased as Adam learned more items. Approximately 150 trials were needed to teach imitation of the use of each of the first three objects, but only 60 trials to teach the sixth object. The same pattern can be observed with regard to other skills. Teaching to follow verbal instructions took progressively fewer trials as number of learned items increased (i.e., 277 to 30 trials, except for the first item). Teaching of gross motor imitation took 1 to 100 trials, with the first item being the easiest to teach. Teaching of the second item took 100 trials and then gradually fewer, with only 25 trials needed to teach the last item (see Figure 6). Hence, Adam’s progress showed a pattern of accelerated learning.

Social Validity

Both instructors rated every part of the staff training to be useful or very useful (average score 4.7). In addition, they rated the training as having increased their teaching effectiveness and the skills of the children to a great extent (average score 5). Both instructors also agreed that the discrete trial teaching approach had increased the children’s attention span considerably (average score 5). In addition, both instructors rated the teaching technique as more effective than other, previously tried techniques (average score 5) and that it would be useful to them in teaching other children as well (average score 5).

Discussion

The goal of this study was to measure the effects of staff training in behavioral teaching techniques on skills of children with developmental disabilities and to assess the generalization of both staff and children’s skills to new settings and tasks. Results indicate that a short workshop and a few weeks staff training can increase the effectiveness of teaching to a considerable extent and thereby greatly improve the skills of children with developmental disabilities. Moreover, skills of instructors and children generalized to a group setting and the instructors were able to teach new tasks using written instructions without further training. These results are generally in accordance with previous findings and are an addition to the few studies on behavioral staff training of preschool personnel working with children with developmental disabilities.

Effects of staff training on teaching. Baseline measures revealed that instructors’ teaching had little in common with discrete trials teaching. Only a small percentage of S^P's was presented in a way that makes correct responding more likely (Anderson et al., 1996; Koegel et al., 1977; Blanc & Ruggles, 1982). Often the instructions lacked clarity or consistency with previously presented instructions, and many times the child was not attending to the task or teacher. In addition, consequences were rarely used in a deliberate manner to increase correct responding; often they were omitted or they were not in accordance with the child’s response. Hence, many learning opportunities were lost during baseline.

Prompts were more often correctly applied during baseline than S^P's and consequences. Prompting usually occurred simultaneously with the S^P and was generally sufficient to help the child respond correctly. Before intervention, prompts were frequently used, making independent correct responding rare. In part, this could be due to too advanced tasks being targeted for instruction during baseline, making independent correct responding nearly impossible. Tasks selected for instruction during staff training fitted the children’s skill level better. Training emphasized building rapid and frequent responding in order to make generalization more likely and facilitate learning of more complex skills (Johnson & Layng, 1994).

There were almost no correctly performed learn units in the instructors’ teaching during baseline. Only one learn unit occurred across nine teaching probes, which approximates a rate of 0.02 units per minute on average during baseline. Interestingly, there was virtually no differences in the rate of learn units in each instructors’ teaching, despite the difference in their education and experience. This low rate of learn units during traditional teaching is in accordance with previous research. For example, teachers in Albers and Greer’s (1991) study used 0.41 learn units per minute during baseline.

Staff training had a considerable effect on the instructors’ use of behavioral teaching techniques. Correct discrete trials teaching increased from 16-31% during baseline to 92-
95% during intervention phases. These results are consistent with previous research on training teachers in discrete trials teaching (e.g., Arco & Millett, 1996; Koegel et al., 1977). Instructors reached the performance criteria for accurate discrete trial teaching after three to seven training sessions. Taking into account the 5-hour workshop and approximately one hour per training session, the instructors needed 8 to 12 hours of training to be able to apply the teaching technique adequately. This number of training sessions is roughly equivalent to the amount in Arco and Millett’s study, where instructors received seven to fourteen 20-minute training sessions in addition to 9 hours of instruction and practice. These 11 to 14 hours of training enabled the teachers to teach new tasks with help of written directions and minimal feedback only (Arco & Millett).

The instructors’ teaching accuracy remained high throughout the study, even at the beginning of teaching new tasks. There were clear signs of transfer of training effects in Hanna’s teaching; continually fewer training sessions were needed to reach mastery when teaching a new task. One of the admission criteria for instructor participants in this study was interest in learning a new teaching technique, and both instructors turned out to be enthusiastic about improving their teaching effectiveness. It is not clear whether the same positive results would have been achieved without this criterion of admission.

Concurrent with improved teaching accuracy the average rate of learn units increased from 0–0.03 per minute during baseline probes to 2.3–2.4 per minute during the experimental phases. This abrupt change in rate of learn units as a result of behavioral staff training is in accordance with previous research (e.g., Albers & Greer, 1991; Ingham & Greer, 1992; Selinske et al., 1991). In Albers and Greer’s study, for example, the rate of learn units tripled when teachers were encouraged to ask more questions of students and present consequences for their students’ responding.

As soon as the instructors started using the behavioral teaching techniques, correct responding on behalf of the children increased. This positive effect of staff training on client’s skills has also been found in previous research (Albers & Greer, 1991; Ingham & Greer, 1992; Kissel et al., 1983; Koegel et al., 1977; Selinske et al., 1991; Thorisdottir, 1993).

Effects of staff training on client skills in this study were greater than in Smith et al.’s (1992) study where a 1-week workshop in behavioral theory and treatment techniques did not have any measurable effect on group home client functioning. Unlike Smith et al.’s, study, staff training in this study took place in the workplace, enabling the instructors to practice teaching skills with their actual students. In addition, measures of children’s skills in the current study only included those targeted during staff training, not other areas of functioning.

Although the staff training had a general positive effect on the children’s skills there were a few complications. David showed considerable defiance, which slowed down the teaching process. Differential reinforcement of on-task behavior was not successful since eliminating reinforcers (for noncompliance and throwing objects) in and outside the teaching environment proved impossible (e.g., because of lack of parental involvement). Due to these behavior difficulties, items involving manipulation of objects were removed from David’s imitation tasks.

Another problem that arose in teaching David, concerned unexpected difficulties labeling objects receptively. Although David quickly learned to point to an object on request, he had great difficulty discriminating between objects. Despite diverse prompting strategies (position, sign language, voice inflection, and physical prompting), attempts to teach David receptive labeling of items proved unsuccessful. Many factors could have contributed to this poor outcome, such as too rapid fading of prompts and/or too swift changes in prompting approaches due to time limitations, inaccuracies in S^b presentation, lack of proficiency in more advanced teaching skills or David’s lack of necessary prerequisite skills.

Generalization of skills from a one-to-one teaching setting to a group setting. The new learned skills of both instructors and children generalized to a group setting. Both instructors continued to use discrete trials accurately while teaching the child participants in a group setting, but Dora taught the whole group with choral responding while Hanna allocated dif-
fferent tasks to the other children. Both approaches worked well, although making sure that all the children responded correctly proved to be time-consuming, resulting in a lower rate of learn units for David than in the one-to-one setting.

This generally small decrease in the rate of learn units when teaching was moved to a group setting is much less than in Kamps, Walker, Maher, and Rotholz (1992) where the rate of trials decreased from 3.5 trials per minute to 1.5 per minute when students were transitioned from a one-to-one setting to small groups. The groups in that study, however, differed from the groups in the present study in that group members were older (5 to 21 years old) and all had been diagnosed with autism and/or other developmental disabili-

ties.

The children’s skills generalized very well to the group settings. Their percentage of correct responding was as high and even higher than in the one-to-one setting. This good performance in a group setting is in accordance with Kamps’ et al. (1992) study where elementary students maintained attending and learning skills following transition from a one-to-one to a small group setting. In another study (Taubman et al., 2001), preschoolers with autism and other developmental disabilities were also successfully taught several skills with a group discrete trial teaching approach.

Taubman et al. (2001) results challenge the widely held tenet that individual discrete trial teaching is a necessary prerequisite for teaching in a group setting (Handleman et al., 1991; Kamps et al., 1992; Lovaas, 1996). However, the group discrete trial teaching approach “while demonstrated to be effective, represents a complicated instructional meth-
odology” (Taubman et al., p. 217) and seems to require experienced teachers and thorough training to be achievable. Thus, further research is required to determine exactly what training is required to enable instructors to use discrete trials to teach new skills in group settings. Given the little knowledge and expe-

rience of the instructors in the present study it was considered advisable to start out in a one-
to-one setting.

Results of this study show that after children with developmental disabilities have reached adequate performance levels in a given skill through discrete trial teaching, correct responding can be maintained in a group setting. Research (e.g., Kamps et al., 1992) shows that when taught in a group, skills take longer to teach. However, considering other advantages of teaching in a group setting, such as the potential of increased social interactions, one-to-one teaching should be restricted to the beginning of teaching new tasks.

Generalization of instructors’ skills to teaching new tasks. During the consultation phase of the study, the instructors’ teaching performance did not drop when starting to teach a new task. This indicates some generalization of skills across teaching tasks, although it is hard to rule out the effects of guidance. Formal measures of generalization of teaching skills to new tasks, 1 and 4 months after training ended, showed that instructors were able to teach new skills by following only written directions.

These positive results on generalization of skills to teaching of new tasks are coherent with findings in Koegel et al.’s (1977) study, where training led to generalized use of behavior modification procedures, and Arco and Millett’s (1996) findings, where training enabled instructors to teach new tasks with minimal feedback. One factor these two studies have in common with the present one is that instructors were trained until they showed at least 80% correct overall performance across several consecutive sessions. These findings are important for the development of effective staff training procedures.

Other studies have shown limited generalization to new tasks (Smith et al., 1992; Thorisdottir, 1993). In Smith et al. staff trainees’ application of one-to-one teaching proce-
du res reached only 39% correct on average in programs taught in the workshop and 30% correct in generalization programs, although their performance was statistically better than a control group receiving no training. After 1 week of training, there was no evidence of any beneficial effects on clients in the group homes. Other factors, such as differences in setting and clients, probably also contributed to this lack of generalization.

This study has several limitations. A multiple baseline design across only two instructors allowed just one replication of the effects of workshop and training on teaching skills. In
addition, the instructors in this study were highly motivated to acquire a new approach to
teaching. Thus, it is not known whether the
findings apply to preschool personnel in gen-
eral. Moreover, baseline data were not col-
clected in the group setting, thus making the
effects of training in the one-to-one setting on
teaching in the group setting unclear.

Taken together the results of this study are
generally in accordance with previous findings
and add to the scarce research literature on
staff training of preschool personnel working
with children with developmental disabilities.
This study did not explore ways to teach new
skills in a group setting and although recent
findings (Taubman et al., 2001) suggest that
discrete trials can be used to teach new skills
in a group setting, necessary prerequisite skills
on behalf of instructors remain to be deter-
mimed. Future research should focus on estab-
lishing viable approaches to train preschool
personnel to effectively teach new skills in an
integrated setting. In addition, it is important
to determine the effects of different training
variables used in this study as well as the opti-
mal rate, sequence or duration of certain staff
behaviors necessary or sufficient to teach chil-
dren effectively (Vollmer, Roane, Ringdahl, &
Marcus, 1999). More research is needed in
this area to establish appropriate criteria for
training in order to make staff training more
precise and efficient.

References
Albers, A. E., & Greer, R. D. (1991). Is the three-
term contingency trial a predictor of effective
354.
Teaching new skills to young children with au-
tism. In C. Maurice, G. Green, & S.C. Luce (Eds.),
Behavioral intervention for young children with autism:
A manual for parents and professionals (pp. 181–
194). Austin, TX: Pro-Ed.
Arco, L., & Millett, R. (1996). Maintaining instruc-
tional behavior after on-the-job training with pro-
cess-based feedback. Behavior Modification, 20,
300–320.
Bayley, N. (1993). Bayley Scales of Infant Develop-
ment: Second Edition. San Antonio, TX: The Psychologi-
cal Corporation.
NJ: Prentice Hall.
Crow, R., & Snyder, P. (1998). Organizational be-
havior management in early intervention: Status
and implications for research and development.
In Dennis H. Reid (Ed.), Organizational Behavior
Management and Developmental Disabilities Services:
Accomplishments and Future Directions (pp. 131–156).
training in special education settings. Education
and Training in Mental Retardation, 22, 205–217.
Intensive behavioral treatment at school for 4-to
7-year old children with autism. Behavior Modifica-
tion, 26, 49–68.
Gardner III, D. M. Sainato, J. O. Cooper, T. E.
Heron, W. L. Heward, J. E. Shelle, et al. (Eds.),
Behavior analysis in education: Focus on measurably
superior instruction (pp. 161–171). Pacific Grove,
Greer, R. D., McCorkle, N., & Williams, G. (1989). A
sustained analysis of the behaviors of schooling.
Behavioral Residential Treatment, 4, 113–141.
Greer, R. D. & McDonough, S. H. (1999). Is the
learn unit a fundamental measure of pedagogy?
The Behavior Analyst, 22, 5–16.
Handleman, J.S., Harris, S.L., Kristoff, B., Fuentes,
for preschool children with autism. Language,
Speech, and Hearing Services in Schools, 22, 107–110.
Harchik, A. E., Sherman, J. A., Hopkins, B. L.,
Stouse, M. C., & Sheldon, J. B. (1989). Use of
behavioral techniques by paraprofessional staff: A
review and proposal. Behavioral Residential Treat-
ment, 4, 331–357.
student and teacher responses in observed and
generalized settings as a function of supervisor
observations. Journal of Applied Behavior Analysis,
25, 153–164.
Jahr, E. (1998). Current issues in staff training. Re-
search in Developmental Disabilities, 19, 73–87.
Johnson, K. R., & Layng, T. V. J. (1994). The Morn-
ingside model of generative instruction. In R.
Gardner III, D. M. Sainato, J. O. Cooper, T. E.
Heron, W. L. Heward, J. E. Shelle, et al. (Eds.),
Behavior analysis in education: Focus on measurably
superior instruction (pp. 173–197). Pacific Grove,
Kamps, D., Walker, D., Maher, J., & Rotholz, D.
(1992). Academic and environmental effects of
small group arrangements in classrooms for stu-
dents with autism and other developmental dis-
abilities. Journal of Autism and Developmental Disor-
ders, 22, 277–299.
An institutional staff training and self-manage-
ment program for developing multiple self-care


Received: 2 May 2005
Initial Acceptance: 17 June 2005
Final Acceptance: 22 August 2005
High School Teachers’ Perceptions of School-to-Work Transition Practices in Taiwan

Meng-chi Chan and Janis G. Chadsey
University of Illinois at Urbana-Champaign

Abstract: The purpose of the study was to investigate practices that high school teachers in Taiwan use to facilitate school-to-work transition, and to determine the importance ratings of each of these transition practices. A questionnaire with 28 transition practices was developed and sent out to the 125 participants in 24 special high schools in Taiwan; 106 completed surveys were used for data analysis. Results from this study identified several key transition practices that were important and also revealed several transition practices that were less valued or less implemented by Taiwanese teachers. Limitations and implications are discussed to improve future research, practice, and cultural diversity in teacher education and transition practices.

Inspired by federal legislation (e.g., IDEA and its 1997 Amendments), practitioners and researchers have made many efforts to improve transition outcomes for students with disabilities. Several follow-up studies of former special education students, however, have revealed negative postsecondary transition outcomes. For example, The NCD (National Council on Disability, 2004) analyzed research on transition, postsecondary outcomes for 14 to 22 year old youth and young adults with disabilities for the past three decades. They reported that 27% of students receiving special education graduated from high school with diplomas while 75% of their peers without disabilities graduated with diplomas. Furthermore, over one third of students with disabilities dropped out or discontinued their education for unknown reasons. The low graduation rate of students with disabilities indicated there was less likelihood of them being employed, receiving postsecondary education, and living independently.

Previous research in U.S. has indicated the importance of identifying critical transition practices in order to improve the quality of transition services and student outcomes (Hughes et al., 1997). Several studies have tried to identify important components and skills essential to specific areas associated with successful transition outcomes (Agran, Snow, & Swaner, 1999; Benz, Johnson, Mikkelsen, & Lindstrom, 1995; Foley & Mundschenk, 1997; Kerka, 2000; Miner & Bates, 1997; Zhang, Katsiyannis, & Zhang, 2002). However, even with these studies, little consensus exists about identification of best transition practices (Johnson & Rusch, 1993; Kohler, 1993). Hughes et al. and Rusch (1992) also suggested that service delivery would be more coordinated and transition outcomes improved if best practices could be identified and disseminated to practitioners.

Efforts have been made to identify comprehensive school-to-work (STW) transition practices in U.S. (Aspel, Bettis, Test, & Wood, 1998; Hughes et al., 1997; Kohler, 1993; Kohler, 1998; Kohler, DeStefano, Wermuth, Grayson, & McGinty, 1994; Mahan & Baer, 2001; Zhang et al., 2002). Practices identified have generally been related to areas focused on student planning and development, vocational education, interagency/interdisciplinary collaboration, family involvement, and program structures and policies. However, few empirical findings have confirmed that the implementation of these practices lead to successful transition outcomes (Kohler, 1993). In addition, many transition practices lack social validation from practitioners (Blanchett,
2001; Knott & Asselin, 1999; Wolfe, Boone, & Blanchett, 1998). The importance of studying teacher perceptions to improve teacher practices has also been noted by researchers in general education (Pajares, 1992; Richardson, 1996). Studies are needed to understand practitioners’ views of best transition practices because what teachers believe are important may have a great impact on what and how they deliver transition services.

In Taiwan, transition is a still new concept. Infused by U.S. literature and legislation, the first regulations related to the provision of transition services, the 1998 Regulation Rules of Special Education Act, mandated that transition services in Taiwan should be identified in the Individualized Education Programs for students from kindergarten to 12th grade. However, the dissemination of research on best practices in transition did not seem to influence many teachers in Taiwan, and most special educators thought transition only meant referral to agencies or job placement (Chen, 2002). Instead of addressing transition issues in a more comprehensive way, most of the transition studies in Taiwan seemed to focus primarily on investigating vocational education (Chou, Yeh, & Chan, 2003; Lin & Shih, 2003) or interagency and interdisciplinary collaboration (Lin, 2004). Chen and Chang (2003) studied transition services needs and services received by youth with disabilities in Taiwan. They found that transition services needed by youth with disabilities were greater than services they received. Chen and Chang suggested that teachers might have had delivered quality transition services if they had more knowledge and skills about transition services delivery.

In order to improve transition service delivery and teacher preparation, it is important to know what transition practices special education teachers in Taiwan implement and what practices they believe are important. Although this information would benefit programs in Taiwan, knowledge of transition practices used and valued by teachers in Taiwan might also benefit programs in the United States. Several researchers have called for multicultural/linguistic teacher preparation in special education (Geenen, Powers, & Lopez-Vasquez, 2001; Obiakor, 2001; Voltz, 1998) and the recruitment and retention of culturally and linguistically diverse teachers (Campbell-Whatley, 2003; Dillard, 1994; Patton, Williams, Floyd, & Cobb, 2003). Transition practices from Taiwan could contribute to the understanding of culture diversity in relation to best transition practices, and may lead to the improvement of teacher preparation programs in the United States by providing teachers with information that can improve their sensitivity to cultural differences.

In Taiwan, most secondary-aged students with disabilities who receive transition services are enrolled in special high schools or special education classes in vocational high schools. To understand current transition practices in Taiwan and teachers’ perceptions of the importance of those practices, this study investigated Taiwanese teachers’ perceptions to transition practices identified in the U.S. There were two reasons for using U.S. transition practices in this study: (a) there is a lack of a comprehensive and rich literature about important transition practices in Taiwan, and (b) it was believed important to understand how well these practices could (or could not) fit into the Taiwanese culture and education system. Past research has indicated that the primary disability of students might be a differentiating factor for students’ transition needs and services (Chen & Zhang, 2003), and training received might influence teachers’ perceptions of transition related issues. Therefore, this study also examined if type of training and primary disability of students taught would have an impact on teachers’ perceptions of transition practices. The following research questions were addressed:

What transition practices do special high school teachers in Taiwan implement and how do special high school teachers rate the importance of these practices?

To what extent do demographic factors (e.g., disability of students, teacher preparation programs) influence special high school teachers’ current transition practices and their importance ratings of transition practices?

Method

Survey Instrument

A questionnaire, Survey of School-To-Work Transition Practices (SSTWTP), was used in this
study. The conceptual framework of the survey was based on the NTA (National Transition Alliance for Youth and Disabilities) Transition Practice Framework (NTA, 1998). Based on the consensus of transition experts and field practitioners, the framework grouped transition practices into five categories (i.e., student-focused transition planning, student-focused development, interagency/interdisciplinary collaboration, family involvement, structures/policies). The SSTWTP used the NTA framework, but combined the first two categories into one category student-focused transition planning and development because of their similar features. Specific questionnaire items for the SSTWTP were primarily adapted from the Nomination packet: Promising transition practices and programs for youth with disabilities (NTA), which was used to evaluate exemplary transition programs in the U.S. The SSTWTP was first developed in English and then translated into Chinese, and each edition was piloted with three practitioners from Taiwan. Based on their comments, the SSTWTP was revised and translated back into English.

The SSTWTP contained two parts: (a) demographic information (i.e., gender, age, years of teaching special education and special high school, certification status, types of certificate program attended, disabilities of students taught), and (b) implementation and importance ratings of STW transition practices. A 3-point Likert scale was used in the implementation ratings (1 = not often implemented, 2 = sometimes implemented, and 3 = very often implemented) and a 4-point scale was used in importance ratings (1 = unimportant, 2 = somewhat important, 3 = important, and 4 = very important). Twenty-eight transition practices were identified and grouped into four categories: (a) student-focused transition planning/development, (b) interagency-interdisciplinary collaboration, (c) family involvement, and (d) structures/policies. The participants were asked to rate how often they implemented each practice and the importance of the practice for improving the transition outcomes of youth with disabilities.

Population and Participants

The population of interest was 866 Taiwanese teachers in 24 special high schools that were established for students with moderate to severe disabilities who wanted to receive high school education but generally did not pass the entrance exams for regular high schools. There were five types of special high schools, based on the primary disabilities of the enrolled students: mental retardation, hearing impairments, visual impairments, physical impairments, and multiple disabilities. A sample of 125 teachers was recruited for this study using a systematic sampling method to randomly select 14.5% of the teachers from each school. The contact person in each school, the head of the Office of Student Practicum Counseling (similar to the vocational coordinator in U.S.), was asked to recruit every fourth person on their teacher lists as a participant until they had 14.5% of their school population. Surveys were returned by 116 teachers for a response rate 92.8%. Excluding surveys with missing data for more than 5 items, 106 surveys were used for data analysis.

Of the 106 participants, 58.5% were female \( n = 62 \) and 41.5% were male \( n = 44 \). The majority of the participants (77.3%) were 31 to 54 years old, and 17% were 30 or under. The majority of teachers, 84.9% \( n = 90 \), had been teaching special education for over 4 years, and 79.2% \( n = 84 \) had been teaching in special high schools for over 4 years. At the time of the study, 67 out of the 106 participants (63.2%) taught students with mental retardation. Regarding certification status, 97 of the participants (91.5%) were certified special education teachers and received their training from the following type of certificate programs listed in most-to-least order of training intensity: (a) 4-year college certificate program \( n = 17 \), (b) 2-year post-college certificate program \( n = 22 \), similar to the Master certificate programs in U.S. but without a degree, (c) master/40-credit program \( n = 26 \), which is more research-oriented, and (d) 20-credit certificate program \( n = 19 \), one-semester curriculum for certified general educators.

Data Collection and Analysis

Data collection was completed over a 6-week period of time. First, a formal letter was sent to the contact person in each school to inform them about the study. One week after the
formal letter was sent out, the investigator made telephone calls to the contact persons in order to ask for their help in participant recruitment and data collection. Surveys, along with the instructions for recruitment and data collection procedures, were mailed to contact persons once they consented to facilitate the study. Contact persons were asked to follow the recruitment procedures to select participants and to distribute and collect the surveys. After questionnaires were distributed, the investigator tracked the status of the surveys with the contact persons once a week. Once all the completed surveys were collected, the contact persons mailed them back to the investigator using a self-addressed envelope.

For data analysis, frequency and percentage were used for descriptions of the demographic information and the implementation and importance ratings. In addition, standard deviations and means were also provided for the rating scales. Analysis of variance (ANOVA) was used to test if the following variables had statistically significant influences on teachers’ implementation and importance ratings of transition practices: (a) whether the primary disability of the students taught was mental retardation, and (b) the type of teacher preparation programs. Since the majority of the participants in this study were teachers of students with mental retardation, and only a few teachers taught students with other primary disabilities, this study focused on examining the group differences between teachers of students with or without mental retardation. With regard to differences in type of teacher preparation programs, the major difference among the four teacher preparation programs was the training intensity. Researchers found that hours and type of training might influence teachers’ perceptions of transition services (Baer, Simmons, & Flexer, 1996; Katsiyannis, deFur, & Conderman, 1998). Therefore, the training intensity of the four different types of teacher training programs was chosen for studying group differences. When a significant group difference was found, Scheffe’ test was used in order to find out which group means were significantly different.

Results

Internal Consistency

Cronbach’s coefficient alpha was used to determine the internal consistency of items in each category of transition practices. Several researchers suggest a minimum reliability of .70 for research purposes (Nunnally, 1978; Siegle, 1997). As shown in Table 1, Cronbach’s alpha scores for the four categories ranged from .69 to .89, and were all above .70 except for the category interagency/interdisciplinary collaboration in implementation ratings ($\alpha = .69$). Although the alpha score was slightly below .70, the items in this category were retained because the literature has identified the importance of interagency and interdisciplinary collaboration. Overall, questionnaire items met the research criteria for internal consistency.

Ratings and Nomination of Transition Practices

Mean rating scores of the transition practices were high, ranging from 2.02 to 2.75 ($SD$ ranged from .50 to .74) in the implementation ratings and from 3.14 to 3.73 ($SD$ ranged from .44 to .76) in the importance ratings. Small standard deviations indicated small variability in participants’ responses. Especially for the importance ratings, over 97% of the ratings were 3 (important) and 4 (very important).

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cronbach’s Alpha Scores by Questionnaire Category</strong></td>
</tr>
<tr>
<td><strong>Category</strong></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Student-Focused Transition Planning/Development</td>
</tr>
<tr>
<td>Interagency/Interdisciplinary Collaboration</td>
</tr>
<tr>
<td>Family Involvement</td>
</tr>
<tr>
<td>Structures and Policies</td>
</tr>
</tbody>
</table>

Transition Practices in Taiwan / 283
Because there were on average only two to three responses in the somewhat important and unimportant categories for each practice, these two categories were included in the important category when presenting results.

Table 2 and Table 3 show descriptive statistics of the 10 transition practices with the highest rated scores. Practices identified in both scales (i.e., implementation & importance) overlapped with one another. In particular, the top four practices in both scales were the same: job placement prior to exit, instruction addresses employment skills, functional/community-referenced curriculum, and community-based work experiences prior to exit. Other overlapping practices included: establish linkage/relationship among school/agents, establish collaborative agreements among schools/service providers, review goal progress annually, and schools support full access/participation in STW activities. Two practices, paid work experiences prior to exit and provide access to postsecondary education, had the lowest importance rating scores ($M = 3.14$). Paid work experiences also had the lowest implementation rating scores ($M = 2.02$).

Table 2 and 3 also show the mean implementation and importance ratings of the four transition categories: student-focused transition planning/development ($M = 2.43$ and 3.51), interagency/interdisciplinary collaboration ($M = 2.34$ and 3.55), family involvement ($M = 2.19$ and 3.37), and structures/policies ($M = 2.24$ and 3.45). Those transition practices identified as being the most implemented and important were primarily from the categories student-focused transition planning/development and interagency/interdisciplinary collaboration. The practices from the category family involvement and structure/policies were less valued and implemented.

**Group Differences**

In both implementation and importance ratings, no statistically significant differences with an alpha level of .05 were found between teachers of students with and without mental retardation in their total and sub-category scores. Additionally, no statistically significant differences ($p < .05$) were found in the implementation ratings among teachers from

**TABLE 2**

Frequency, Percentage, Mean, and Standard Deviation of Implementation Ratings of Transition Practices

<table>
<thead>
<tr>
<th>Category/Practices</th>
<th>Not Often (1)</th>
<th>Sometimes (2)</th>
<th>Very Often (3)</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student-Focused Transition Planning/Development</td>
<td>2.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job placement prior to exit</td>
<td>5  4.7</td>
<td>16 15.1</td>
<td>82 77.4</td>
<td>2.75</td>
<td>.53</td>
</tr>
<tr>
<td>Instruction addresses employment skills</td>
<td>2  1.9</td>
<td>27 25.5</td>
<td>76 71.7</td>
<td>2.70</td>
<td>.50</td>
</tr>
<tr>
<td>Functional/community-referenced curriculum</td>
<td>6  5.7</td>
<td>29 27.4</td>
<td>71 67.0</td>
<td>2.61</td>
<td>.60</td>
</tr>
<tr>
<td>Community-based work experiences prior to exit</td>
<td>6  5.7</td>
<td>31 29.2</td>
<td>69 65.1</td>
<td>2.59</td>
<td>.69</td>
</tr>
<tr>
<td>Review goal progress annually</td>
<td>11 10.4</td>
<td>32 30.2</td>
<td>60 56.6</td>
<td>2.48</td>
<td>.67</td>
</tr>
<tr>
<td>Specified transition service providers</td>
<td>10 9.4</td>
<td>40 37.7</td>
<td>56 52.8</td>
<td>2.43</td>
<td>.66</td>
</tr>
<tr>
<td>Interagency/Interdisciplinary Collaboration</td>
<td>2.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establish linkages/relationship among school/agents</td>
<td>7  6.6</td>
<td>38 35.8</td>
<td>61 57.5</td>
<td>2.51</td>
<td>.62</td>
</tr>
<tr>
<td>Collaborative agreements established among schools/service providers</td>
<td>12 11.3</td>
<td>34 32.1</td>
<td>60 56.6</td>
<td>2.45</td>
<td>.69</td>
</tr>
<tr>
<td>Family Involvement</td>
<td>2.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structures/Policies</td>
<td>2.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schools support full access/participation in STW activities</td>
<td>8  7.5</td>
<td>42 39.6</td>
<td>56 52.8</td>
<td>2.45</td>
<td>.63</td>
</tr>
<tr>
<td>Transition policy/procedures/practices</td>
<td>8  7.5</td>
<td>43 40.6</td>
<td>55 51.9</td>
<td>2.44</td>
<td>.63</td>
</tr>
</tbody>
</table>
different training programs. However, results showed statistically significant differences in importance rating scores among teachers from the four different teacher preparation programs in the total scale, $F(3,80) = 4.19$, $p < .01$, and in the following transition categories: student-focused planning/development, $F(3,80) = 3.29$, $p < .05$; interagency/interdisciplinary collaboration, $F(3,80) = 5.08$, $p < .01$; and structures/policies, $F(3,80) = 2.78$, $p < .05$. Further analysis with the Sheffe’s test revealed only one statistically significant difference between teachers from college certification programs and teachers from 20-credit programs for the category interagency/interdisciplinary collaboration, $F(3,80) = 4.19$, $p < .05$: the mean rating scores of the former were 1.29 larger than that of the latter.

**Discussion**

Results revealed several significant findings. First, transition practices related to family involvement and structures/policies were less often implemented and valued by Taiwanese educators, even though the literature in both the United States and Taiwan have identified the importance of family involvement (Asselin, 1995; Chen, 1997; Katsiyannis et al., 1998; Knott & Asselin, 1999; Lin, 1998) and structures and policies (Baer et al., 1996; Chen, 1997). It is possible that the Taiwanese teachers’ lower ratings for family involvement and structures/policies might be due to their lack of training on these issues. Lower ratings for family involvement may have also been due to findings by Caplan, Hall, Lubin, and Fleming (1997) who found that the degree of parent involvement decreased when their children were older, especially after entering middle or high school. Additionally, the long distance between home and school and low family socioeconomic status might discourage the implementation of these transition services. Most Taiwanese students attending special high schools live far away from school. The long distance between home and school might dis-
encourage parents from being involved in transition planning, which in turn, might effect teachers’ ratings. Clark and Kolstoe (1995) stated that families with varying socioeconomic status showed differences in their willingness and capacity to be involved in their children’s transition planning and curriculum; families with higher socioeconomic status were more engaged in transition planning and were more capable of retrieving information about their children’s welfare. According to the League of Disability (2001), 44.78% of Taiwanese families with children with disabilities were unable to make ends meet, while 46.15% had low to moderate socioeconomic status; lower socioeconomic status might create difficulties for family involvement. Teachers’ beliefs often come from personal experience, experiences with schooling and instruction, and experience with formal knowledge (Richardson, 1996). The lack of knowledge and experiences with families might keep Taiwanese teachers from implementing transition practices related to family involvement.

A second finding was that providing access to postsecondary education was not considered important nor was it implemented frequently by the participants, even though this is a key practice in U.S. literature (Aspel et al., 1998; Kohler et al., 1994; Stodden & Whelley, 2004). According to the Special Education Transmit Net (2004), there are 5,757 students with disabilities receiving postsecondary education and most of them have disabilities other than mental retardation. To receive a college education in Taiwan, one must take the College Entrance Examination (CEE), similar to the SAT in U.S., or be selected through a recommendation system, which means one’s academic performance in one of the core subjects (e.g., math, science, English, Chinese) must be top 5% of the peers in order for being recommended by his/her school. For students who take the CEE, their scores must be higher than the test standard of that year for being assigned to a college. Because of the education system and corresponding policies, it has been difficult for students with disabilities to receive a postsecondary education in Taiwan (Chen, 2000). For most special high school teachers in this study, the majority of their students had mental retardation.

Therefore, it was not surprising that the participants considered postsecondary education unimportant for most of their students; this finding also may reveal a difference in educational philosophy between the United States and Taiwan.

Finally, statistically significant group differences in importance ratings were found among teachers from training programs of different training intensities, but the post hoc comparisons only revealed that teachers from 20-credit programs had significantly higher rating scores for the category “interagency/interdisciplinary collaboration” than teachers from 4-year college certificate programs. Baer et al. (1996) found that hours of training received by teachers had an impact on their implementation of transition practices, and Knott and Asselin (1999) found that teacher implementation and importance ratings were highly correlated. Thus, intensity of training might also have an impact on teacher importance ratings. However, this study indicated that teachers with more intense training did not place higher values on transition practices. A possible explanation for this finding might be that teachers from 20-credit programs used to be general education teachers and their teaching experiences in general education might have impacted their perceptions of the importance of interagency/interdisciplinary collaboration. In addition, teachers might learn their transition knowledge through other sources (e.g., inservice training, professional journals). It might be necessary to identify participants whose transition knowledge were primarily learned from their certificate programs in order to find out if there are any group differences in these teacher preparation programs.

Implications for Practice

This study has several implications for practice. First, teacher preparation and inservice training in Taiwan may need to put more emphases on knowledge and competencies in implementing culturally-sensitive practices related to family involvement and structures/policies. Katsiyannis et al. (1998) found inservice training was an effective way to improve the implementation of transition practices. Thus, improvement of inservice training as
well as preservice preparation programs might result in the enhancement of transition outcomes for students with disabilities. In addition, the Taiwanese government and other stakeholders should also make efforts to establish clear policies and guidelines in transition related issues in order to assist the implementation of transition practices (Chen, 1997).

This study found that Taiwanese teachers placed less value on transition practices related to family involvement. This finding may have implications for U.S. teachers working with students with disabilities from Taiwan. Parents from different cultural backgrounds, such as Taiwan, might have a different level of involvement in their children’s education. Caplan et al. (1997) stated that parents from minority groups had less involvement in their children’s school events and activities for several reasons, such as past bad experiences with schools, deference to education in their cultures, and poor English skills. Teachers prepared in the United States should be aware of and sensitive to cultural diversity and this should be an important consideration when promoting family involvement in transition planning. In addition, it is also necessary to recruit and retain teachers with diverse cultural background and languages in special education in order to meet the needs of increasing culturally and linguistically diverse population with disabilities (Campbell-Whatley, 2003; Patton et al., 2003).

Limitations

There were several limitations to this study that need to be addressed. First, the representativeness of the participants might reflect a bias. Although the contact person in each school was asked to randomly select the participants, it was unknown if the sample was randomly selected or if it was voluntary sampled. In addition, the majority of study participants were teachers of students with mental retardation. Their perceptions may not represent teachers of students with other disabilities, such as learning disabilities. Furthermore, the sample might be too small to generalize the findings to the general population. Second, the implementation and importance rating scores in this study showed low variability. As noted earlier, this might have been caused by a positive response bias. Third, survey research measures the perceptions of the participants, not their actual behaviors. It is unknown if there were discrepancies between what participants perceived and what they really did. Finally, the survey instrument was developed based on the U. S. literature and translated into Chinese. Some terms might have been misunderstood by the participants. Even with these limitations, this study makes a contribution by studying a more global and international view of transition practices and providing insight for practice and future research.

Recommendation for Future Research

After reviewing results and limitations of this study, further investigation is warranted in several areas. First, future studies should recruit teachers of students with other types of disabilities and should use a larger sample in order to improve the comprehensiveness and representativeness of the research sample. Second, teacher knowledge and training could be an important predicator of teachers’ perceptions. It would be important to investigate how well prepared teachers are in the area of transition and to note their perceptions for each transition practice and where they get their knowledge or training. Lastly, interviews with more in-depth questions could be combined with survey research. Qualitative data would be useful to interpret quantitative data, overcome the limitation of survey research, and enhance the research quality.

References


Music Therapy and the Education of Students with Severe Disabilities

Jennifer Stephenson
Macquarie University

Abstract: Music therapists regard music therapy as a valuable intervention for students with moderate to severe intellectual disability or multiple disabilities, but many special educators would regard it as a controversial practice, unsupported by empirical research. This paper reviews the goals and strategies used by music therapists working with students with severe disabilities and the purported outcomes of music therapy. The recent research base that could validate music therapy as an effective educational intervention is reviewed. There is little evidence to support the use of music therapy as an educational intervention, but what evidence there is suggests that music therapy sessions, when planned in collaboration with educators, may provide a context for eliciting and practicing communication skills. Such music therapy interventions should be individually planned and monitored to ensure educational outcomes are achieved.

Music therapy has been broadly defined by a music therapist with an interest in people with disabilities as “...the use of music as a therapeutic tool for restoration, maintenance, and improvement of psychological, mental and physiological health and for the habilitation, rehabilitation, and maintenance of behavioral, developmental, physical and social skills – all within the context of a client-therapist relationship” (Boxill, 1985, p. 5). Music therapy is used with a wide range of populations – people in hospital, people with psychiatric disorders, older people, people in hospices, people with neurological problems, people with autism and adults and children with intellectual disability (Aldridge, 1993).

Music therapy is seen by music therapists as a useful contribution to the education of students with special education needs (Aldridge, Gustorff, & Neugebauer, 1995; Davison & Edwards, 1998; Patterson, 2003). Music therapists view music therapy as a “...well established profession similar to occupational therapy and physical therapy” (American Music Therapy Association, 2002, p. 1). Others, however, would place it in the category of controversial or non-proven approaches in special education, due to the lack of empirical evidence regarding its effectiveness (Dempsey & Foreman, 2001; New York State Health Department, 1999). Its use is not included in standard texts on the education of students with moderate to severe disabilities (see Snell & Brown, 2000; Westling & Fox, 2004) and it is not included in articles reviewing effective approaches to educating this population (Broder & Cooper-Duffy, 2003; Wolery & Schuster, 1997). Music therapy does, however, meet some of the criteria for identifying controversial practices identified by McWilliam (1999) and Herbert, Sharp, and Gaudiano (2002). For example, promotion of its use relies largely on anecdotal and case study evidence, some proponents are hostile to scientific evaluation (Aigen, n.d.), and proponents claim benefits for a very wide range of conditions. Even so, the use of music therapy in educational programs for students with severe disabilities is widespread (Chase, 2004; Ockelford, Welch, & Zimmerman, 2002; Smith & Hairston, 1999; Stephenson, 2004).

How then, should schools and teachers working with students with severe disabilities approach this therapy?

This paper focuses on the use of music therapy in educational settings for school students.
with moderate to profound intellectual disability and multiple disabilities, excluding students with autism spectrum disorders. It will discuss prevalence of use of music therapy in special education settings, describe goals and practice of music therapy with students with severe disabilities, review recent research and then consider possible applications of music therapy in the education of students with severe disabilities.

**Use Of Music Therapy In Special Education Settings**

Music therapy is becoming increasingly associated with special education, particularly with the education of students with severe disabilities. It is seen as a desirable component of education by some parents of students with disabilities (Fidler, Lawson, & Hodapp, 2003). In the U.S. context, it is recognized in educational settings as a “Related Service” that may be provided to assist a child with special education needs (Mattson, 2001; Patterson, 2003). Also in the U.S., 12% of music therapists report working with people with developmental disabilities (Chase, 2004) and increasing numbers are working with children with developmental disabilities in school settings (Chase; Smith & Hairston, 1999). Smith and Hairston, in a survey of American music therapists who worked in schools, found that 78% worked with children with developmental disability and 71% with children with multiple disabilities. In the U.K. a survey found just over a third of schools enrolling students with severe or profound learning difficulties had a music therapist working on site, although the authors suggest as few as 2% of students may have received therapy (Ockelford et al., 2002). Similarly, in Australia, a survey of the websites of schools enrolling pupils with high support needs found just over a quarter of the sites included information about music therapy in the school (Stephenson, 2004).

**Goals Of Music Therapy**

Meadowes (1997) in a review of music therapy for children with severe and profound multiple disabilities, described six goals of music therapy for this population. The first is “fulfilling the child’s basic needs” (p.4) which involves creating a trustworthy and responsive environment. The second is “developing the child’s sense of self” (p. 4) where the child builds relationships with musical instruments, music and the therapist. The third is “establishing or re-establishing interpersonal relationships” (p. 4). The fourth is “developing specific skills” (p. 5) such as eye contact, reaching, or using a switch within musical activities. The fifth is “dispelling pathological behaviour” (p. 5), and the sixth is “developing an awareness and sensitivity to the beauty of music” (p. 5).

Music therapists may approach these goals in a number of ways, depending on their philosophical and theoretical beliefs. Hooper (2002) cited Moranto (1993) as enumerating at least 123 forms of music therapy. It seems that only a subset of these are used with children with severe disabilities. According to Meadowes (1997) music therapists working with students with severe disabilities may focus on music as recreation, as a reinforcer for desired behavior, as a means to develop other skills and knowledge, and/or to “heal.” Similar outcomes for music therapy within special education are described by the American Music Therapy Association (1999a, b). Music therapists using the Nordoff-Robbins approach, often used with people with severe disabilities, emphasize the development of communication and relationships through music (Nordoff-Robbins Music Therapy Australia, n.d.). Daveson and Edwards (1998), writing in the Australian context also noted the use of music therapy in achieving academic goals, in teaching gross and fine motor skills, in developing social interaction skills and in using music as a motivator for other tasks. In a survey of assessment tools used by music therapists in the U.S., Chase (2004) reported that therapists assessed motor skills (fine and gross), communication skills, social skills and cognitive/academic skills. Pellitteri (2000), who also identified speech and language, gross and fine motor skills, academic, behavior, social and aesthetic goals as part of music therapy, sees this crossing of several areas as a strength of music therapy in special education.
Meadowes (1997) described a range of activities that may occur in music therapy sessions, such as the child using musical instruments alone or with the therapist. The therapist may respond to the child’s use of an instrument. The child and therapist may improvise together or the child may learn a specific music skill. The therapist may introduce activities that promote non-music skills, but are related to music and the instruments such as shaking or turn taking. The therapist may encourage the child to vocalize or sing, or work on listening and comprehension. This may be at the level of provision of sensory stimulation or at a higher level. Movement activities may also be included to promote both whole body and fine motor skills through formal movement or by the child spontaneously responding to music. Children may move independently or be assisted by adult helpers. Music and movement sessions may be run with the involvement of a physiotherapist and may be highly structured with specific movement goals or improvised. Different therapists may use the same range of activities in different ways depending on their orientation (Meadowes, 2002).

Pellitteri (2000) described the typical group musical therapy session in a special education setting in the U.S. as commencing with a greeting song, moving into activities such as singing songs, playing instruments individually, in turns or in groups and moving to music in a directed or spontaneous way before a final song to close the session. He notes that these activities may be used by teachers to reinforce the therapy sessions. He does not consider this as music therapy but rather the use of music in education, because he sees the relationship between the therapist and the child as an essential element of music therapy.

Perry (2003) described a similar structure in sessions in Australia. Precomposed and improvised songs were used as well as improvised accompaniments related to children’s responses. Musical instruments and voice were used in turn taking and synchronous interactions.

The diversity of goals and activities of music therapists, lead to a range of claims made about the benefits of music therapy to children with severe disabilities, including the development of communication, social skills and purposeful movement (Aldridge et al., 1995; Duffy & Fuller, 2000). As Davison and Edwards (1998) point out, many of the general goals of music therapists are broadly congruent with the goals of special education. What is lacking is the evidence that students can in fact achieve those goals through music therapy. Although music therapists claim benefits, these claims depend more on anecdotes and descriptive case studies that empirical evidence (see Nordoff & Robbins, 1985). Boxill (1985) made many claims for the benefits of music therapy for people with disabilities, and included vignettes and case studies, but there is little research beyond case study reports to validate the claims made.

Erdonmez (1991) reviewed relevant literature regarding the efficacy of music therapy with different groups. She found that music may affect stereotypical behaviors such as rocking, that children with profound intellectual disability responded more to voice than to musical instruments, and that vocalization was more likely when music was soft and when the sound source was close to the child’s head. She located one study that showed music was a reinforcer for some people with intellectual disability. On the other hand, Green, Reid, Canipe, and Gardner (1991) who assessed the preferences of people with profound multiple disabilities found that neither rock music nor soft music was a preferred stimulus for any of the students they assessed. None of the studies summarized by Erdonmez speak to the efficacy of music therapy in bringing about important and significant change in people with severe disabilities.

In a review of studies on music therapy published in medical journals between 1983 and 1990, Aldridge (1993) mentions one descriptive study on children with multiple disabilities. In a general conclusion on music therapy research up to 1990 he stated (p. 28), “...there is a general absence of valid clinical research material from which substantive conclusions can be drawn.” He also makes the
surprising suggestion that if music therapy is an accepted treatment for children with disabilities, it “requires no further scientific investigation to support its incorporation as part of a treatment plan.” (p. 29). This conclusion would be unacceptable to those who advocate for evidence-based practice.

Even now, ten years on from Aldridge’s (1993) review and call for more scientific investigation, there appear to have been few controlled studies of treatment outcomes, particularly for children with intellectual disability. This has been recognized within the field itself (Aldridge et al., 1995; Duffy & Fuller, 2000). Ockelford et al. (2002) in their discussion of the role of music in the education of students with severe disabilities noted that there is “scant contemporary literature” (p. 178) on this topic, although there is a growing body in relation to music therapy itself. There has always been a focus on more qualitative methods within the field of music therapy, and though these may be ideal for exploring the perceptions of music therapists and the processes of music therapy, they do not provide the quantitative data necessary to demonstrate change in children that can be clearly attributed to music therapy (Aldridge et al.).

Review Of Recent Research On Music Therapy

In order to review more recent work, a search was conducted for studies on the outcomes of music therapy that included children aged 4 to 18 years old with moderate to profound intellectual disability or multiple disabilities (excluding autism spectrum disorders), published between 1995 and 2004. Studies needed to have some educational relevance. Those addressing purely medical aspects were not included. Hand searches of the Australian Journal of Music Therapy, British Journal of Music Therapy, Journal of Music Therapy and Music Therapy Perspectives were carried out. In addition, searches using the electronic data bases PsychINFO, Pro-quest Education Journals and Expanded Academic, were carried out using the search term “music therapy.” The titles and/or abstracts of articles were viewed, or where necessary the whole article, and only those meeting the above criteria were included. The reference lists of all articles located in these searches were also searched for further relevant articles.

This search located only seven studies as summarized in Table 1. Where studies included younger children or adults, and results for these participants are included separately, they are not included in the table.

Aldridge et al. (1995) reported a group study that compared two small groups (5 and 3 children) of children with disabilities who commenced music therapy at different times, and received therapy over a twelve-month period. Effects were assessed using the Griffiths Mental Developmental Scales and a scale developed by Nordoff and Robbins that has never been validated. Results suggested that music therapy had a small positive effect, but other causes could not be ruled out. Designs of this kind are flawed because they do not compare the treatment with a control treatment and gains may have been seen if the children had spent an equivalent amount of time in other non-music activities with responsive adults. It is important to design studies that demonstrate that it is the particular activities in music therapy that promote development. This study is also problematic in that it does not give a clear description of the procedures used so that they could be replicated by other music therapists.

Braithwaite and Sigafoos (1998) compared the effect of musical antecedents and social antecedents on the communicative responsiveness of five pre-school children with severe disabilities. For three of these children the musical antecedents appeared to be slightly more effective in eliciting use of existing communication skills, although there was some overlap in the range of results. Further work needs to be done to explore which factors were responsible. For example, students may have been more motivated to request a musical instrument in the music condition than a book in the social condition, or the musical conditions may have provided a general motivational boost to some children. This study does show, however, that small n designs (in this case an ABAB design) can demonstrate empirically the effects of music therapy on particular behaviors of individual children.

Duffy and Fuller (2000) explored effectiveness of a music therapy program for improving social skills in children with moderate in-
intellectual disability. Two programs were devised to teach the same skills, one with pre-recorded music for musical activities and the other with substitute activities for the musical activities. Staff was trained in the implementation of the programs with matched groups of children at four centers. Social skills were assessed pre and post program. Results showed

**TABLE 1**

Music Therapy Studies 1995–2004

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Design</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aldridge et al.</td>
<td>Twelve children aged 4 to 6.5 yrs, developmental age 1.5 to 3.5 yrs.</td>
<td>Two groups, first group received three months individual therapy while second group received none, then second group received therapy while first group rested. Replicated once.</td>
<td>Four children lost to study. Larger mean changes to Griffiths Quotient during therapy than resting, mostly in hearing and speech subscale.</td>
</tr>
<tr>
<td>Braithwaite &amp; Sigafoos</td>
<td>Five children aged 3.5 to 4.5 years with severe delays in adaptive behaviour and language development.</td>
<td>Compared effects on communicative responsiveness of two antecedent conditions, social interaction and social interaction plus music using an ABAB design.</td>
<td>Moderate increases in responsiveness for three children in the music plus social condition.</td>
</tr>
<tr>
<td>Duffy &amp; Fuller</td>
<td>32 children, aged 5 to 10 years with moderate intellectual disability.</td>
<td>Compared music therapy social skills teaching with non-music social skills teaching.</td>
<td>The music therapy intervention was not more beneficial than the non-music approach. None of the three conditions were more effective than baseline in eliciting and maintaining alert behavior.</td>
</tr>
<tr>
<td>Ghetti (2002)</td>
<td>Six children aged 7 to 17 with profound disabilities.</td>
<td>Within subjects design compared behavior state in baseline (talking to child) with behavior state in three different musical conditions (rhythmic stimulation, song singing, instrument playing).</td>
<td>None of the three conditions were more effective than baseline in eliciting and maintaining alert behavior.</td>
</tr>
<tr>
<td>Perry (2003)</td>
<td>Ten students aged 5 to 11 years with severe and multiple disabilities.</td>
<td>Qualitative study describing communication in music therapy sessions.</td>
<td>Musical interaction can provide a context for communicative interaction.</td>
</tr>
<tr>
<td>Yashuhara &amp; Sugiyama</td>
<td>Three girls aged 4, 5 and 6 with Rett Syndrome.</td>
<td>Received 40, 40 and 12 sessions of music therapy. No details provided. Were receiving other intervention as well.</td>
<td>No detail on how change assessed. Some improvements noted, not sustained in one child.</td>
</tr>
</tbody>
</table>
that both groups improved their social skills, and that the music therapy intervention was not more beneficial. However these results may be open to debate as the music was pre-recorded and therefore does not reflect the responsive, improvisational approach taken by many therapists.

Ghetti (2002) described a study that explored the effect of various musical conditions (rhythmic stimulation from the therapist playing a drum, song singing intended to elicit vocalization, and playing rhythm instruments) on the behavioral state of students with profound disabilities. She found none of the musical treatments were better than baseline conditions, where the therapist simply talked to the students, at maintaining students in an alert state. However, for all sessions the students were wheeled from their classroom to a new environment, and all sessions included interaction with the therapist and therefore baseline levels of alert state may have been higher than levels in a typical classroom environment. Much more detailed individualized analysis of behavior state may be necessary to capture any effects of music therapy on individual children.

Studies by Hill (1997) and Yasuhara and Sugiyama (2001) both concerned girls with Rett Syndrome. Both provided little detail of the music therapy practices, or of how improvements were assessed. Neither show that improvements were due to music therapy and not to normal development or other interventions.

The study by Perry (2003) provided a qualitative description of music therapy sessions and children’s communication skills. It did not aim to demonstrate effects and will be discussed later in this paper.

These more recent studies add little to the earlier research. It seems the more carefully a study is designed, the less likely it is that significant positive effects will be demonstrated. Many of these studies show that the use of music may provide a positive context and have a motivational effect for some children, but these uses would not require a music therapist. There are studies outside the music therapy field that demonstrate these effects of music clearly. For example, a study by Durand and Mapstone (1998) clearly showed that for two adults and one child with intellectual disability, challenging behavior was much more likely when they listened to slow beat music than when they listened to fast beat music. Negative facial expressions were also more common during slow beat music. Music thus seems to moderate the conditions which would normally produce problem behavior in some people. This study clearly identified, through functional analysis, the conditions that lead to challenging behavior by the participants and then through further individualized analysis clearly demonstrated the impact of music during the 10 minute assessment sessions. This study suggests that appropriate music may help produce a generally “positive context” (p. 376). Indeed the teachers of one participant in this study successfully introduced the use of fast beat music during difficult tasks to promote participation. A similar finding in relation to music was made by Carey and Halle (2002) who after carrying out functional analysis of the self-injurious behavior (SIB) of a boy with severe intellectual disability, found that the behavior was maintained by escape from task demand and by access to music. The student was then found to have lower rates of SIB when music was available during demand conditions. Although these studies illustrate potential positive effects of the use of music as a reinforcer, use of music in this way would not appear to require a music therapist.

There is then, still very little evidence to show that musical therapy can result in the achievement of important educational outcomes for students with severe disabilities. There is a clear need for additional good research in this area which is methodologically sound and which incorporates clear criteria for demonstrating effects.

Discussion

Approaches To The Use Of Music Therapy In Schools

Special education has not been immune from the adoption of fads and unproven controversial therapies, such as facilitated communication (Huebner & Emery, 1998) and similarly the allied health fields have generated a number of unproven practices which are directed at students with special education needs such
as sensory integration and auditory integration (McWilliam, 1999; Shaw, 2002). Should music therapy be regarded as a fad treatment, or should it be taken more seriously? The research base for the use of music therapy with students with severe disabilities is certainly sparse, and educators would be justified in viewing its use with considerable suspicion.

Given that music therapy is widely used in schools, how should educators approach its use? Procedures described by Brunk and Coleman (2000) in the U.S. for determining whether or not music therapy should be included as a related service in a student’s individual educational program (IEP) may be a starting point for making decisions about who might benefit educationally from participation in music therapy sessions.

In the U.S., since music therapy can be an allied service and incorporated into a child’s IEP, it must contribute to the achievement of the educational goals set in the IEP. Ideally the music therapist would work with the child’s family, teachers and other professionals to carry out assessment, to develop the goals and objectives and teaching programs. However music therapists may also provide consultation or provide a written recommendation describing their services to the IEP team. Music therapy can then be legally included if an appropriate assessment is carried out that demonstrates to the parents and to the school district that music therapy would help the child attain a measurable annual goal, progress in the general curriculum, participate in extracurricular activities and/or participate in activities with children without disabilities. (Brunk & Coleman, 2000; Mattson, 2001; Patterson, 2003).

Brunk and Coleman (2000), both music therapists, detail their recommended process for assessment (SEMTAP) by a music therapist as part of the IEP process in U.S.. They suggest that music therapy may be appropriate for students who have demonstrated that they can be motivated to attempt or complete tasks by music, who use additional communication modalities, who have shown interest in music or musical instruments and who retain information from songs. If music therapy appears to be appropriate, they recommend that specific individualized assessment (best carried out by a music therapist who will not be contracted to deliver the service) should then focus on specific IEP goals that could be addressed within music therapy sessions. The student’s current performance on these goals with and without music therapy procedures should be compared. The aim of this detailed assessment is to clearly demonstrate that the student is helped to perform activities directed at achievement of IEP goals by music therapy strategies, and to indicate whether the music therapist should provide direct services in pull out sessions or work within the classroom.

In the sample assessment provided, music therapists presented quite formal tasks such as matching colors to printed color names, using graphic symbols to choose, using scissors to cut and ordering three pictures in sequence. The therapist also observed the child’s responses to music and its possible role as a motivator. This form of individualized assessment seems to have been widely adopted by music therapists working with children with disabilities in the U.S.. A survey of assessment strategies used by music therapists found that 70% reported using the SEMTAP (Chase, 2004). It certainly provides a structured process to determine whether or not music therapy would be of educational benefit to an individual student.

**Potential Benefits Of Music Therapy**

For students with more severe and multiple disabilities, an area of possible benefit of music therapy may be in the area of communication skills. Communication skills have been a particular focus of attention for music therapists, with many case study reports and anecdotes claiming improvement in this area (see for example Aldridge et al., 1995; Boxill, 1985; Nordoff & Robbins, 1985). Many music therapists clearly appreciate the importance of developing intentional and presymbolic communication for students with severe disabilities and the ways these can be elicited in interaction with a responsive partner (Boxill; Aldridge et al.; Perry, 2003; Wigram, 1999). It appears that music therapists who emphasize the building of relationships through musical activities are particularly aware of early communication skills, particularly non-verbal communication, and the ways in which children...
with communication delays or disorders might be encouraged or motivated to communicate within musical activities. Such therapists describe the use of improvisations, which are responsive to the child to promote eye contact, turn taking, sharing, joint attention and other early communication skills (Boxill; Perry; Voigt, 1999; Wigram).

Perry (2003), in a qualitative observational study of students with severe and multiple disabilities in music therapy sessions highlighted the potential of these sessions to develop early communication skills in these children through opportunities for joint attention, turn taking, initiating, responding and sustaining attention and the like. She provided a discussion of the ways in which music therapists may interact with children, which mirror the responsive interactions between parent and child believed to support communication development (O’Kane & Goldbart, 1998). From her observations, it appears that music therapists can provide a range of communication opportunities and encourage children to be responsive to these opportunities, but they face difficulties with students who are not alert and who appear unmotivated. Similarly, Braithwaite and Sigafoos (1998) demonstrated increased use of existing communication skills in a music therapy context. Despite such research however, it appears that more research in this area to examine acquisition and use of early communication skills within a music therapy context as one form of a responsive environment to elicit these skills.

**Recommendations**

The most fruitful approach to use of music therapy as an educational intervention in schools for students with severe disabilities may be to focus on its use as context for teaching and practicing early communication skills. In the absence of a research base and clear guidelines for practice, music therapists, speech pathologists and educators should work together to carefully assess existing communication skills and the conditions under which they may be elicited and reinforced (see for example O’Kane & Goldbart, 1998). However, it is debatable whether the use of music and musical activities as a responsive context for teaching requires the use of a music therapist. Special educators have traditionally embedded the teaching of functional skills in motivating activities and routines, and the use of musical activities in this way may exploit the motivating and reinforcing effects of music in conjunction with pedagogical practices that have a sound research base.

It is likely that interventions employing music therapy would need to be tailored to individual responses, as not all students will respond to music and not all are likely to respond in the same way (Braithwaite & Sigafoos, 1998; Green et al., 1991). Those working with students should have clear goals in mind, recorded as observable outcome statements and progress towards the outcomes should be assessed across all contexts, including music therapy sessions. There is a clear need for further research to demonstrate the effectiveness of music therapy for achieving a range of goals in the education of students with severe disabilities. Given the lack of empirical support for educational benefits arising from music therapy, it may be appropriate for schools to reassess the rationale for music therapy programs and to determine whether or not the program is providing anything beyond a pleasant experience and an enjoyable context for teaching skills. If music therapy is used with students with severe disabilities, it should be used responsibly after appropriate assessment, with clear aims and with ongoing monitoring to demonstrate that learning is occurring.

**References**


Ockelford, A., Welch, G., & Zimmerman, S. (2002). Music education for pupils with severe or profound and multiple difficulties – current provi-


Received: 10 March 2005
Initial Acceptance: 1 May 2005
Final Acceptance: 15 August 2005
Reliability and Validity of the Pervasive Developmental Disorders Rating Scale and the Gilliam Autism Rating Scale

Ronald C. Eaves and Suzanne Woods-Groves
Auburn University

Thomas O. Williams, Jr. and Anna-Maria Fall
Virginia Polytechnic Institute and State University

Abstract: The psychometric properties of the Pervasive Developmental Disorders Rating scale (Eaves, 2003) and the Gilliam Autism Rating Scale (Gilliam, 1995) were investigated in this study. One hundred thirty-four individuals with autism, other pervasive developmental disorders, or conditions frequently confused with autism participated in the study. The results indicated that, with one exception, the reliability of the scores from both instruments met or exceeded standards for use in screening decisions. The reliability of the total scores from both instruments exceeded .90. Validity coefficients computed between the two sets of scores indicated that the instruments measured similar constructs (e.g., $r_{pddrs \text{ total } \times \text{ gars \ total}} = .84$). The scores from both instruments discriminated between children with autism and children who were not autistic to a statistically significant degree.

The purpose of this research was to examine the reliability and validity of two screening instruments: the Gilliam Autism Rating Scale (GARS; Gilliam, 1995), and the Pervasive Developmental Disorder Rating Scale (PDDRS; Eaves, 2003). The GARS is purported to identify individuals with autistic disorder, one of five pervasive developmental disorders (PDD) defined in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000). The PDDRS purports to identify individuals with PDD.

In an effort to estimate the reliability and validity of the GARS and PDDRS, we employed four sets of analyses. First, Salvia and Ysseldyke (2004) have established critical reliability values for specific decisions. For making eligibility and classification decisions, these authors recommended a minimum reliability coefficient of .90. For screening decisions, they recommended a minimum reliability coefficient of .80. Although the GARS and PDDRS are not recommended for eligibility decisions by their respective authors, they can reasonably be held to a standard of $r \geq .80$.

A standard way of estimating the validity of an instrument is to compare the correlations of its scores with another instrument designed to serve the same, or a similar purpose. Because the GARS and PDDRS do purport to serve similar purposes, we established as a second objective of our research to compute validity coefficients between sets of scores obtained from the GARS and the PDDRS on the same participants.

Although it does not provide compelling evidence, it is reasonable to expect such instruments as the GARS and PDDRS to discriminate between autistic-PDD groups and non-autistic-non PDD groups. It was our third objective to test this reasonable expectation. While positive results support the validity of instrument, a more severe test is the determination of whether or not the instrument successfully classifies individuals. Our fourth set of analyses sought to assess the classification accuracy of the GARS and PDDRS for individuals. This included the estimation of sensitivity and specificity for the two instruments on a sample comprised of participants with autistic disorder, Asperger’s disorder, pervasive developmental disorder-not otherwise specified,

Correspondence concerning this article should be addressed to Ronald C. Eaves, Department of Rehabilitation and Special Education, 1228 Haley Center, Auburn University, AL 36849. Email: eaves11@charter.net
and participants with disabilities who were also suspected of having a PDD. We considered these analyses to pose the most severe test of the instruments. The following questions were addressed in this study:

1. To what extent do the GARS and PDDRS measure their respective dimensions accurately?
2. To what extent do the GARS and PDDRS measure the same constructs?
3. Do the GARS and PDDRS discriminate between groups of individuals with different diagnoses?
4. To what extent do the GARS and PDDRS classify individuals with different diagnoses accurately?

Method

Participants

In this study 66 participants rated 134 individuals either diagnosed with PDD (i.e., autistic disorder (n = 86), Asperger’s disorder (n = 11), pervasive developmental disorder-not otherwise specified (n = 15)], or some other disability that is often confused with PDD (n = 23). The second group included one child with cerebral palsy, four children with developmental delays, two children with mild mental retardation, seven individuals with moderate mental retardation, four individuals with multiple disabilities, one youngster with severe-profound mental retardation, and four children with severe communication disorders. Although we did not record the number, several of these participants (e.g., the child with cerebral palsy) were selected for assessment specifically because they were thought to have autism or some other PDD. The participants resided in one of five southeastern states or Washington, D.C. Teachers of children with pervasive developmental disorders, college teaching interns, and parents and guardians participated in the study. Ninety-seven of the PDDRS and GARS ratings were completed by teachers (72.39%), nine ratings were completed by graduate interns (6.72%), and 28 ratings were completed by parents and guardians (20.90%). The mean length of time that the rater had known the child was 2.82 years (SD = 4.17). Signed informed-consent documents were obtained from the parents or legal guardians of the children rated.

The raters reported the participants’ formal labels and asserted that they were being served according to those labels. In Alabama, where most of the participants resided, autism is defined as, “a developmental disability that significantly affects verbal and nonverbal communication and social interaction evident before age three that adversely affects educational performance. Other characteristics often associated with autism are engagement in repetitive activities and stereotyped movements, resistance to environmental change or changes in daily routines, and unusual responses to sensory experiences. The term does not apply to children who have an emotional disturbance” [Alabama Administrative Code, 2004, 290-8-9.03(1)(a)]. The diagnosis of autism is commonly determined by a team of individuals consisting of medical, clinical, psychiatric, psychological, and/or other qualified personnel trained in the area of autism assessment.

Of 134 participants, 17.16% (n = 23) were female and 82.84% (n = 111) were male. The ethnicity of two participants was unknown. Of the remaining participants, 59.85% were white (n = 79), and 40.15% were African-American (n = 53). The participants ranged in age from 3-to-26 years, with a mean of 9 years, 8 months (SD = 4 years, 7 months). The socioeconomic status (SES) of the participants was estimated using scores based on the occupation of the head of household (U.S. Bureau of the Census, 1963). Scores can range from 1 (undefined personal services) to 99 (physicians). The midrange SES score (50) is assigned to such occupations as assistant librarians, bakers, and bricklayers. The mean SES of the sample was 71.10 (SD = 24.13; range = 99), indicating that the sample was generally of middle class, but exhibited a high degree of variability.

Instruments

Pervasive Developmental Disorder Rating Scale. The PDDRS is a rating scale developed by Eaves (1990; Eaves & Hooper, 1987–1988). It contains 51 items that measure three dimensions: Arousal, Affect, and Cognition. The items were developed following an examina-
tion of the classic literature on autistic disorder (e.g., Kanner, 1943; Lovaas, Freitag, Gold, & Kassorla, 1965; Rimland, 1964) and a summation of behavioral characteristics of PDD drawn from the DSM-III-R (American Psychiatric Association, 1987), research literature, existing instruments, and the clinic files of individuals with autistic disorder and PDD. Raters are requested to evaluate each item independently using a five-point Likert scale according to the degree to which the individual exhibits the behavior described. The PDDRS was normed on 814 individuals diagnosed with pervasive developmental disorders. Raw scores may be transformed into standard scores (M = 100, SD = 15) and percentile ranks.

The internal consistency of the PDDRS was estimated using the split-half technique followed by a Spearman-Brown adjustment for scale length (Eaves, 2003). The reliability coefficients were as follows: (a) r_{pddrs total} = .92, (b) r_{arousal} = .90, (c) r_{affect} = .84, and (d) r_{cognition} = .79. Test-retest reliability was estimated with two samples. In the first sample, reliability was based on pairs of ratings collected over a mean interval of 8.33 months from the same 18 raters. The reliability coefficients were r_{pddrs total} = .91, r_{arousal} = .89, r_{affect} = .87, and r_{cognition} = .87. The second sample reflected both test-retest and interrater reliability inasmuch as two different raters completed PDDRSs on 80 participants over a relatively long test-retest interval of 14.20 months. The reliability coefficients were much lower for this sample: r_{pddrs total} = .48, r_{arousal} = .53, r_{affect} = .40, r_{cognition} = .44.

The reliability of the PDDRS was also examined with a sample of 567 individuals labeled with some variant of PDD (Williams & Eaves, 2002). The participants were divided into two groups based on chronological age (CA). The low-CA group was made up of 456 individuals ranging in age from 1-to-12 years and the high-CA group ranged in age from 13-to-24 years. Alpha coefficients for the low-CA group ranged from .75 to .89, with a Total Score coefficient of .89. Alpha coefficients for the high-CA group ranged from .77 to .89 for the three scales, with a Total Score coefficient of .89.

The test-retest reliability of the PDDRS was examined with a sample of 40 individuals who had been rated twice by the same rater (Williams & Eaves, 2002). The mean interval between ratings was 9.50 months (SD = 2.96; range = 24). Coefficients for test-retest reliability ranged from .86 to .92 for the three scales, with a Total Score reliability of .92. The results of the reliability studies indicated that the internal consistency and stability of the PDDRS were adequate for research purposes, met or exceeded the minimum requirements for screening purposes, and were stable over time for both the individual being rated and the rater.

The criterion-related validity of the PDDRS and the Autism Behavior Checklist (ABC; Krug, Arick, & Almond, 1993) was examined by comparing data for both instruments with a sample of 107 children known to be diagnosed with autism and 32 children who were diagnosed with disabilities frequently confused with autism (Eaves, Campbell, & Chambers, 2000). Results for the total scores for the PDDRS and the ABC showed that the instruments measured similar constructs (r = .80). Both instruments also significantly discriminated between participants with autism disorder and participants with disorders frequently confused with autistic disorder. The PDDRS had a classification accuracy rate of 88% and the ABC had an accuracy rate of 80%. The PDDRS and the ABC agreed in their classifications for 85% of the 139 participants.

The construct validity of the PDDRS was originally based on the factor analysis of 500 sets of ratings on children with pervasive developmental disorders (Eaves, 1990). Four hundred and thirty-six of the children were diagnosed with autistic disorder. Following a first- and second-order factor analysis of the data, the instrument was reduced to three factors: Arousal, Affect, and Cognition. It was proposed that these factors corresponded to functions associated with the reticular activating system, limbic system, and the cerebrum (Eaves, 1990, 2003; Eaves & Awadh, 1998).

Using a sample of 199 children with autism from 1 to 6 years of age, Eaves and Williams (2006) conducted exploratory and confirmatory factor analyses of PDDRS scores. In the exploratory factor analyses, the three-factor solution best fit the data when compared to one- and two-factor solutions. In the confirmatory factor analyses, the hypothesized second-
order model (i.e., autism was comprised of arousal, affect, and cognition) provided the best fit indices when compared to five competing models. Williams and Eaves (2005) found similar results using a sample of 168 older youngsters with autism.

**Gilliam Autism Rating Scale.** The GARS was designed to assess individuals, ages 3 to 22 years, for autism. Parents, teachers, and other professionals complete it. The GARS consists of 56 items divided into four scales: (a) Stereotyped Behaviors, (b) Communication, (c) Social Interaction, and (d) Developmental Disturbances. Each scale is comprised of 14 items that are said to be indicative of autistic disorder. Respondents rate the frequency of each behavior on a 4-point scale: (a) never observed, (b) seldom observed, (c) sometimes observed, and (d) frequently observed. Each scale raw score is converted into a standard observed, and (d) frequently observed. Each behavior on a 4-point scale: (a) never observed, (b) seldom observed, (c) sometimes observed, and (d) frequently observed. Each scale raw score is converted into a standard score (\(M = 10, SD = 3\)). The scale standard scores are summed and converted into an Autism Quotient (\(M = 100, SD = 15\)). The Autism Quotient is intended to determine the likelihood that a subject has an autistic disorder. It is also used to estimate the severity of the disorder (Gilliam, 1995).

The GARS manual described the Autism Quotient as being comprised of seven categories, ranging from very low to very high probability of autism. Higher Autism Quotients indicate an increased probability of autism. For example, an Autism Quotient of 90 to 100 indicates that the child is probably autistic (Gilliam, 1995). The Autism Quotient may be calculated from two, three, or four scales. Users of the GARS are instructed to use fewer than the four scales in two instances: (a) if the child is nonverbal and does not communicate with others, then the Communication scale is not used; and (b) if the informant is not aware of the child’s developmental history, then the Developmental Disturbances scale is not completed.

Gilliam (1995) described the GARS norm group as consisting of 1,092 children from across the United States and Canada reported to be autistic by parents or teachers. The norms were based on the entire reference sample and were not categorized by gender or age.

The GARS examiner’s manual reported the following estimates for the GARS internal consistency by employing Cronbach’s (1951) coefficient alpha. Reliability estimates for the scores were: (a) Stereotyped Behaviors (\(r = .90\)), (b) Communication (\(r = .89\)), (c) Social Interaction (\(r = .93\)), (d) Developmental Disturbances (\(r = .88\)) and (e) Autism Quotient (\(r = .96\)).

Gilliam (1995) examined the interrater reliability of the GARS. Thirty-five teachers and 79 parents rated 57 participants (43 males and 17 females). The participants had the following diagnoses: autism (\(n = 43\)), mental retardation (\(n = 9\)), emotional disturbance (\(n = 2\)), and multihandicapped (\(n = 3\)). The mean age of the participants was 10 years. Three sets of correlations were computed: (a) teacher-teacher (\(r = .91\)), (b) parent-parent (\(r = .72\)), and (c) teacher-parent (\(r = .95\)) (Gilliam, 1995). By including participants with diagnostic characteristics that are quite different than autistic disorder (i.e., emotional disturbance) Gilliam extended the range of the scores and interrater reliability was predictably inflated (Thorndike, 1982).

The GARS’ test items were derived from the DSM-IV (American Psychiatric Association, 1994) in an effort to ensure the content validity of the instrument. Gilliam (1995) used two item-discrimination criteria to select the final items for the GARS. First, the point-biserial correlations had to be statistically significant at or beyond the .05 level. Second, half of the point-biserial correlations were required to attain or exceed .35 in magnitude. The following median point-biserial correlations were obtained: (a) Stereotyped Behaviors, \(r = .61\); (b) Communication, \(r = .65\); (c) Social Interaction, \(r = .69\); and (d) Developmental Disturbances, \(r = .61\).

Gilliam (1995) compared the GARS with the Autism Behavior Checklist (ABC), a component of the Autism Screening Instrument for Educational Planning (Krug et al., 1993). Sixty-nine participants, randomly chosen from the normative sample, were employed. Forty-nine of the subjects were reported to be autistic while 20 were youngsters with: (a) mental retardation (\(n = 7\)), (b) emotional disturbance (\(n = 7\)), and multiple disabilities (\(n = 6\)). A correlation of .94 was reported for the comparison between the GARS Autism Quotient and the ABC Total.

South et al. (2002) examined the validity of
the GARS by comparing it with the Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 1994), the Vineland Scales of Adaptive Behavior, Survey Form (Sparrow, Balla & Cicchetti, 1984), and the Autism Diagnostic Observation Schedule-Generic (Lord et al., 2000). They found the GARS underestimated the likelihood that the children with autism in the sample would be classified as having autism. A sensitivity of .48 was found. Because there were no non autistic participants in the sample, specificity and overall classification accuracy could not be estimated.

Procedure

Teachers were asked to submit informed-consent documents to the parents or guardians of each child in their classrooms. The informed-consent document described the PDDRS, the GARS, and the nature of the research. During this process 23 parents and five guardians indicated their interest in completing a PDDRS and GARS response forms on their children. Upon receipt of the informed-consent document, PDDRS and GARS response forms were disseminated, completed by the raters, and collected by the first author. For the PDDRS, response forms were scored twice, using Macintosh and IBM computer software (PDDRS Assistant; Eaves, 2005); printouts with matching scores were considered to be accurate. To ensure accuracy each GARS response form was scored twice using the appropriate norms tables in the GARS manual (Gilliam, 1995). The analyses were completed using SPSS 11.0 for Windows (2001).

Results

Table 1 displays the means and standard deviations for the GARS and PDDRS scores. For both instruments the sample standard score means approximated the means for their respective normative samples (i.e., either 100 or 10). Among the observed standard deviations, the GARS Autism Quotient standard deviation (i.e., 19.26) was considerably larger than the normative standard deviation of 15 points.

The first question addressed in this research was, “To what extent do the GARS and PDDRS measure their respective dimensions accurately?” To answer this question, Cronbach’s alpha coefficients were calculated for all PDDRS and GARS scores. Table 1 presents these statistics. The reliabilities of the total scores of both instruments exceeded the cut off for making eligibility-classification decisions (i.e., .90; Salvia & Ysseldyke, 2004). With the exception of the GARS Developmental Disturbances scale, the scores of the remaining scales of both instruments exceeded the commonly cited cut off for screening decisions (i.e., $r = .80$; Salvia & Ysseldyke).

<p>| TABLE 1 |</p>
<table>
<thead>
<tr>
<th>Means, Standard Deviations, and Coefficients Alpha for Gillian Autism Rating Scale and Pervasive Developmental Disorders Rating Scale (PDDRS) Standard Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dimension</strong></td>
</tr>
<tr>
<td><strong>Gillian Autism Rating Scale</strong></td>
</tr>
<tr>
<td>Autism Quotient</td>
</tr>
<tr>
<td>Stereotyped Behavior</td>
</tr>
<tr>
<td>Communication</td>
</tr>
<tr>
<td>Social Interaction</td>
</tr>
<tr>
<td>Developmental Disturbances</td>
</tr>
<tr>
<td><strong>Pervasive Developmental Disorders Rating Scale</strong></td>
</tr>
<tr>
<td>PDDRS Total</td>
</tr>
<tr>
<td>Arousal</td>
</tr>
<tr>
<td>Affect</td>
</tr>
<tr>
<td>Cognition</td>
</tr>
</tbody>
</table>
The second research question sought to determine the extent to which the GARS and PDDRS measure similar constructs. Validity coefficients between the GARS and PDDRS scores are displayed in Table 2. The correlation between the total scores was .84, which indicates a high degree of shared variance between the two instruments. The validity coefficients ranged from .09 to .84 (median \( r = .64 \)). Nominally, three pairs of PDDRS and GARS scores appeared to measure similar constructs: (a) the PDDRS Arousal and GARS Stereotyped Behaviors scores, (b) the PDDRS Affect and GARS Social Interaction scores, and (c) the PDDRS Cognition and GARS Communication scores. The PDDRS Arousal and GARS Stereotyped Behaviors validity coefficient was .84. The PDDRS Affect and GARS Social Interaction validity coefficient was .76. The PDDRS Cognition and GARS Communication validity coefficient was .64.

The third research question asked whether or not the GARS and PDDRS discriminate between groups of individuals with different diagnoses. The results for the analyses of variance for autistic and non-autistic groups and the GARS and PDDRS scores are presented in Table 3. All comparisons were statistically significant. The effect size, as estimated by partial eta squared (\( \eta^2 \)), was .19 for the GARS Autism Quotient and .25 for the PDDRS Total. The results for the analyses of variance for pervasive developmental disorders and non-pervasive developmental disorders groups and the GARS and PDDRS scores are presented in Table 4. All comparisons were statistically significant with the exception of the GARS Developmental Disturbances (\( F(1,80) = 3.28, \ p = .07 \)) and the PDDRS Cognition (\( F(1,132) = 6.43, \ p = .01 \)). The effect size (\( \eta^2 \)) was .12 for the GARS Autism Quotient and .14 for the PDDRS Total. For both sets of analyses, Dunn’s (1961) tables were used to adjust the alpha across multiple comparisons to maintain a constant alpha of .05.

The fourth research question asked, “To what extent do the GARS and PDDRS classify individuals with different diagnoses accurately?” Four analyses were conducted to answer this question. First, two conventional classification accuracy analyses were conducted in which GARS and PDDRS classifications were compared to the participants’ clinical diagnoses. Table 5 displays the results of these analyses. In the first analysis each participant was classified as autistic \(( n = 86 \) ) or not autistic \(( n = 48 \) ); participants with Asperger’s disorder, PDD-not otherwise specified, and other

### Table 2

<table>
<thead>
<tr>
<th>Dimension</th>
<th>GARS</th>
<th>PDDRS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AQ</td>
<td>SB</td>
</tr>
<tr>
<td>GARS Stereotyped Behavior (SB)</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>Communication (Comm)</td>
<td>a 82</td>
<td>a 58</td>
</tr>
<tr>
<td>Social Interaction (SI)</td>
<td>.91</td>
<td>.78</td>
</tr>
<tr>
<td>Developmental Disturbances (DD)</td>
<td>b 73</td>
<td>b 43</td>
</tr>
<tr>
<td>PDDRS Total</td>
<td>.84</td>
<td>.77</td>
</tr>
<tr>
<td>Arousal (AR)</td>
<td>.83</td>
<td>.84</td>
</tr>
<tr>
<td>Affect (AF)</td>
<td>.73</td>
<td>.65</td>
</tr>
<tr>
<td>Cognition</td>
<td>.31</td>
<td>.15</td>
</tr>
<tr>
<td>( n = )</td>
<td>134</td>
<td>134</td>
</tr>
</tbody>
</table>

*Note.* AQ = Autism Quotient. Validity coefficients are in boldface.

\( a\) \( n = 117 \).

\( b\) \( n = 82 \).

\( c\) \( n = 75 \).
(e.g., moderate mental retardation, severe communication disorder) were considered not autistic. In the second analysis each participant was classified as PDD or not PDD; thus, participants with autistic disorder, Asperger’s disorder, and PDD-not otherwise specified were considered PDD (n = 111). The remaining participants were considered not PDD (n = 23).

Although the normal GARS cut off for an autism-non autism decision is an Autism Quotient of 90 (South et al., 2002), in this sample a standard-score cut off of 85 faired at least as well. Using the Autism Quotient of 85 as the criterion, the GARS produced sensitivity, specificity, and overall accuracy estimates of 87.21%, 47.92%, and 73.13%, respectively. The author of the PDDRS recommended that

### TABLE 3

**Analysis of Variance for Diagnostic Label (Autistic-Not Autistic) and the Gilliam Autism Rating Scale (GARS) and Pervasive Developmental Disorders Rating Scale (PDDRS) Scores**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Autistic M (SD)</th>
<th>Not Autistic M (SD)</th>
<th>df</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARS Autism Quotient</td>
<td>103.79 (17.36)</td>
<td>86.54 (17.62)</td>
<td>1,132</td>
<td>30.11</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Stereotyped Behavior</td>
<td>10.63 (3.43)</td>
<td>8.12 (3.25)</td>
<td>1,132</td>
<td>17.01</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Communication</td>
<td>10.79 (3.35)</td>
<td>7.95 (3.03)</td>
<td>1,115</td>
<td>20.65</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Social Interaction</td>
<td>10.37 (3.14)</td>
<td>7.12 (3.07)</td>
<td>1,132</td>
<td>33.46</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Develop Disturb</td>
<td>10.46 (2.81)</td>
<td>8.38 (3.48)</td>
<td>1,80</td>
<td>8.91</td>
<td>.0038</td>
</tr>
<tr>
<td>PDDRS Total</td>
<td>108.44 (14.54)</td>
<td>91.90 (12.64)</td>
<td>1,132</td>
<td>43.68</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDDRS Arousal</td>
<td>107.51 (15.79)</td>
<td>91.02 (14.73)</td>
<td>1,132</td>
<td>35.25</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDDRS Affect</td>
<td>106.74 (14.94)</td>
<td>92.65 (12.80)</td>
<td>1,132</td>
<td>30.30</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDDRS Cognition</td>
<td>106.29 (15.94)</td>
<td>98.04 (14.31)</td>
<td>1,132</td>
<td>8.86</td>
<td>.0035</td>
</tr>
</tbody>
</table>


### TABLE 4

**Analysis of Variance for Diagnostic Label [Pervasive Developmental Disorder-Not Pervasive Developmental Disorder (PDD)] and the Gilliam Autism Rating Scale (GARS) and Pervasive Developmental Disorders Rating Scale (PDDRS) Scores**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>PDD M (SD)</th>
<th>Not PDD M (SD)</th>
<th>df</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARS Autism Quotient</td>
<td>101.11 (18.06)</td>
<td>80.74 (15.94)</td>
<td>1,132</td>
<td>25.17</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Stereotyped Behavior</td>
<td>10.27 (3.40)</td>
<td>7.13 (3.22)</td>
<td>1,132</td>
<td>16.50</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Communication</td>
<td>10.34 (3.38)</td>
<td>7.32 (2.97)</td>
<td>1,115</td>
<td>14.88</td>
<td>.0002</td>
</tr>
<tr>
<td>GARS Social Interaction</td>
<td>9.91 (3.15)</td>
<td>5.83 (3.01)</td>
<td>1,132</td>
<td>32.55</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>GARS Develop Disturb</td>
<td>9.94 (3.20)</td>
<td>8.39 (3.24)</td>
<td>1,80</td>
<td>3.28</td>
<td>.0739</td>
</tr>
<tr>
<td>PDDRS Total</td>
<td>105.53 (14.79)</td>
<td>87.96 (13.44)</td>
<td>1,132</td>
<td>27.71</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDDRS Arousal</td>
<td>104.71 (16.31)</td>
<td>86.61 (13.90)</td>
<td>1,132</td>
<td>24.60</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDDRS Affect</td>
<td>104.10 (14.84)</td>
<td>90.09 (14.81)</td>
<td>1,132</td>
<td>16.99</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PDDRS Cognition</td>
<td>104.88 (16.06)</td>
<td>95.87 (12.45)</td>
<td>1,132</td>
<td>6.43</td>
<td>.0124</td>
</tr>
</tbody>
</table>

individuals obtain standard scores 85 on both the Arousal score and the Total score. Using these criteria, the PDDRS also exhibited somewhat better classification accuracy for autism/non-autism decisions when compared to a cut off of 90; in this analysis sensitivity was 93.02%, specificity was 47.92%, and overall classification accuracy was 76.87%.

When estimating PDD-non PDD classification accuracy, the GARS produced better results using the Autism Quotient criterion of 85. In this analysis sensitivity was 83.04%, specificity was 68.18%, and overall classification accuracy was 80.60%. When contrasted with standard scores of 90 as the cut off for the PDDRS, standard scores of 85 for Arousal and Total scores produced better classification accuracy. For the PDDRS, sensitivity was 83.72%, specificity was 58.33%, and overall classification accuracy was 74.63%. Although the results across criteria were very similar for the GARS and PDDRS, in terms of absolute values, the PDDRS accuracy estimates equaled or exceeded the GARS estimates for 11 of 12 comparisons (mean difference = 2.93%). That is, the PDDRS accuracy estimates exceeded the GARS for eight comparisons, accuracy estimates were identical for three comparisons, and the GARS accuracy estimates exceeded the PDDRS for one comparison.

The third analysis that was conducted to answer the fourth research question investigated the degree to which the GARS and PDDRS agreed with one another on the proper classification of the participants. The GARS and PDDRS agreed that 96 of 134 participants would appropriately be labeled as autistic disorder/PDD. The GARS and PDDRS agreed on the non autistic-non PDD label for 25 of the 134 participants. Thus, the two instruments agreed in their classifications for 121 participants (90.50%) and disagreed on just 13 participants (9.70%).

The last analysis used to answer the fourth research question involved the computation of the phi coefficient (Siegel & Castellan, 1988). The phi coefficient is a measure of the extent of association between two sets of attributes measured on a nominal scale, each of which may take on only one of two values (e.g., autism-non autism or PDD-non PDD). When the phi coefficient was used to estimate the degree of the association between the GARS and PDDRS nominal classifications, the correlation was high and statistically significant (Φ = .74, p = .000).

Discussion

This research investigated the reliability and validity of the GARS and PDDRS. The results generally supported the two instruments for use as screening devices for autistic and other pervasive developmental disorders. The sample means of both instruments were close to their respective normative values of 100 or 10 (depending upon the dimension measured). Although the sample standard deviation of the GARS AQ (SD = 19.26) was excessively large, the remaining standard deviations for
both instruments were reasonably close to the expected values of 15 or 3.

The analysis of the internal consistency of the GARS and PDDRS supported their use as screening devices. Several dimensions provided reliability estimates above .90 (i.e., GARS AQ, GARS Social Interaction, PDDRS Total, and PDDRS Arousal). Only one dimension, the GARS Developmental Disturbances, produced scores with a reliability coefficient below .80 (i.e., \( r = .74 \)). With the exception of the GARS Developmental Disturbances dimension, which was previously reported to have a coefficient alpha of .88 (Gilliam, 1995), the remaining estimates were very similar to those reported in the previous literature.

The concurrent validity evidence produced in this study strongly supported the assertion that the GARS and PDDRS measure similar constructs. For instance, the validity coefficient calculated between the two total scores was .84. Among the other pairs of scores, three matches were found which had the following validity coefficients: (a) GARS Stereotyped Behavior and PDDRS Arousal \( (r = .84) \), (b) GARS Social Interaction and PDDRS Affect \( (r = .76) \), and GARS Communication and PDDRS Cognition \( (r = .64) \). Thus, it may be asserted that the two instruments rank order examinees in much the same way.

Whether the GARS and PDDRS were used to screen individuals with autistic disorder or pervasive developmental disorders, they did discriminate between groups of individuals with different diagnoses in this investigation. Across 18 comparisons of means, only two fell short of statistical significance at the .05 alpha level: GARS Developmental Disturbances and PDDRS Cognition. Both occurred in the PDD-non PDD comparisons.

The fourth research question asked, “To what extent do the GARS and PDDRS classify individuals with different diagnoses accurately?” In our classification accuracy analysis, we used standard-score cut offs of 85 (as recommended for the PDDRS) and 90 (as recommended for the GARS). Although the results were somewhat mixed, both instruments produced better overall classification accuracy when the standard-score cut off of 85 was used. Whether the classifications were based on autism-non autism or PDD-non PDD, the accuracy of the PDDRS either equaled or exceeded the accuracy of the GARS in 11 of 12 comparisons.

Although the overall classification accuracy estimates computed in this study for the GARS \( (M = 75.75\%) \) and PDDRS \( (M = 78.36\%) \) were satisfactory, they were lower than previously published estimates (GARS = 90%, PDDRS = 88.00%). Given that several participants in this investigation were actually suspected of having some form of PDD, we examined the extent to which the GARS and PDDRS agreed with one another in their classifications. First, a cross tabs analysis showed that the two instruments agreed that 96 participants in the sample were autistic-PDD and that 25 participants were not autistic-not PDD. Disagreements regarding the proper diagnosis were found for only 13 participants. Thus, the GARS and the PDDRS agreed on 90.30% (121 + 134) of the participants. Second, the phi coefficient \( (\Phi = .74) \), which estimated the degree of the association between the GARS and PDDRS nominal classifications, indicated a high degree of relationship between the two instruments.

References


Reliability and Validity - PDDRS and GARS / 309


Received: 27 April 2005
Initial Acceptance: 21 June 2005
Final Acceptance: 15 September 2005